1. ΕΙΣΑΓΩΓΗ 2

1.1 Τι ειναι Monte Carlo 2
1.2 Που εφαρμοζεται 2
1.3 Ποια ειναι η προελευση της 2
1.4 Γιατι την χρησιμοποιουμε 2
1.5 Ποια ειναι η απλουστερη μορφη της Monte Carlo 3
1.6 Ποια ειναι τα τρια κυριοτερα χαρακτηριστικα των μεθοδων MC 4
1.7 Πως προσπαθουμε να βελτιωσουμε την μεθοδο 4
1.8 Περισσοτερα για το παραδειγμα Επιτυχιας η Αστοχιας 5

2. ΣΤΟΙΧΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ ΣΤΗΝ MONTE CARLO 6

2.1 Νομος των Μεγαλων αριθμων 6
2.2 Η κατανομη Gauss και ο κανονας των 3 αποκλισεων 6
2.3 Κεντρικο Οριακο Θεωρημα 7
2.4. Το γενικο σχημα των μεθοδων Monte Carlo 9

3. ΟΛΟΚΛΗΡΩΣΗ MONTE CARLO 10

3.1 Η μεθοδος του υπολογισμου 10
3.2 Το σχημα της ολοκληρωσης οταν παρουμε ομοιομορφη τ.μ. 11
3.3 Το σφαλμα της ολοκληρωσης Monte Carlo 12
3.4 Επεκταση στην πολυδιαστατη περιπτωση 14
3.5 Συγκριση με μεθοδους που βασιζονται σε πλεγμα (grid based methods) 14

4. ΠΑΡΑΓΩΓΗ ΤΥΧΑΙΩΝ ΑΡΙΘΜΩΝ 15

4.1 Μεθοδος Μετασχηματισμου 15
4.2 Μεθοδος Αποδοχης - Απορριψης 17

5. ΜΕΙΩΣΗ ΔΙΑΣΠΟΡΑΣ 19

5.1 Αντιθετες μεταβλητες (Antithetic Variables) 20
5.2 Μεταβλητες Ελεγχου (Control Variables) 21
5.3 Matching Moments Μεθοδος 21
5.4 Διαστρωματωμενη δειγματοληψια (Stratified Sampling) 23
5.5 Σημαντικη Δειγματοληψια (Importance Sampling) 23
Ενα Αριθμητικό Παράδειγμα 25
Εκτιμήσεις Σφαλμάτων 27
5.6 Προσαρμοσιμες Monte Carlo – Μεθοδος Vegas 28

6. QUASI-MONTE CARLO ΜΈΘΟΔΟΙ 29

6.1 Monte Carlo enantion quasi-Monte Carlo 29
6.2 Ασυμφωνια (discrepancy) 30
6.3 Koksma-Hlawka ανισοτητα 31
6.4 Μεσο Σφαλμα Ολοκληρωσης 33
6.5 Βασικες Γεννητριες quasi-τυχαιων αριθμων 34
6.6 Περιορισμοι 37

7. ΠΑΡΑΔΕΙΓΜΑΤΑ 42

8. ΒΙΒΛΙΟΓΡΑΦΙΑ 50

1. Εισαγωγή

1.1 Τι είναι Monte Carlo

Η έκφραση *μέθοδος Monte Carlo* είναι πολύ γενική. Περιλαμβάνει κυρίως στοχαστικές διαδικασίες, εκείνες δηλαδή που βασίζονται στην χρήση των τυχαίων αριθμών και της στατιστικής για την λύση προβλημάτων. Γενικά, η μέθοδος Monte Carlo είναι μια αριθμητική μέθοδος για την επίλυση μαθηματικών προβλημάτων μέσω προσομείωσης τυχαίων αριθμών. Χρησιμοποιείται δε στην προσομείωση (simulation) και στην ολοκλήρωση (integration). Μιλώντας γενικά, κάθε πείραμα στο οποίο χρησιμοποιούνται τυχαίοι αριθμοί για την εξέταση του προβλήματος, λέγεται Monte Carlo πείραμα.

1.2 Που εφαρμόζεται

Μέθοδοι Monte Carlo εφαρμόζονται σε πάρα πολλούς επιστημονικούς τομείς, από την οικονομία εώς την πυρηνική φυσική και την χημεία και ακόμη ως τη ρύθμιση της κυκλοφορίας.

1.3 Ποια είναι η προέλευσή της

Η μέθοδος "γεννήθηκε" το 1949, σε ένα άρθρο των N. Metropolis & S. Ulam με τίτλο "Η μέθοδος Monte Carlo" στο Journal of the American Statistics Association. Παρόλα αυτά η θεωρητική βάση της μεθόδου ήταν γνωστή πριν από το 1949, αφού αρκετά προβλήματα στατιστικής λύνονταν μέσω τυχαίας δειγματοληψίας, που είναι στην ουσία η μέθοδος Monte Carlo. Λόγω του γεγονότος ότι η προσομείωση τυχαίων μεταβλητών είναι μια δύσκολη διαδικασία για να γίνει χειρωνακτικά, η γενική αριθμητική χρήση της μεθόδου έγινε πρακτική μόνο με την εμφάνιση και εξέλιξη των υπολογιστών.
Η μέθοδος δανείστηκε το όνομα της πόλης του Πριγκηπάτου του Monaco, που είναι γνωστό για τα καζίνο του, γιατί μια από τις απλούστερες συσκευές παραγωγής τυχαίων αριθμών είναι η ρουλέτα.

1.4 Γιατί την χρησιμοποιούμε

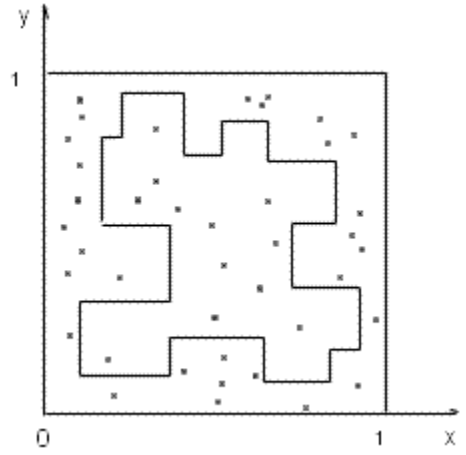
Η χρήση των μεθόδων Monte Carlo στη μοντελοποίηση φυσικών προβλημάτων μας επιτρέπει να εξετάσουμε πολύπλοκα συστήματα που αλλιώς θα ήταν από δύσκολο εώς αδύνατο. Η επίλυση εξισώσεων που περιγράφουν την αλληλεπίδραση δύο ατόμων είναι σχετικά εύκολη…Η λύση όμως των ίδιων εξισώσεων για εκατοντάδες ή χιλιάδες άτομα είναι αδύνατη. Με τις μεθόδους Monte Carlo, ένα μεγάλο σύστημα μπορεί να δειγματιστεί σε έναν αριθμό τυχαίων ρυθμίσεων, και αυτά τα δεδομένα μπορούν να χρησιμοποιηθούν για να περιγράψουμε το σύστημα σαν σύνολο.

1.5 Ποιά είναι η απλούστερη μορφή της Monte Carlo

Η ολοκλήρωση με τη Επιτυχία ή Αστοχία (Hit or Miss) είναι το απλούστερο είδος Monte Carlo.

Παράδειγμα Επιτυχίας ή Αστοχίας:
Ας υποθέσουμε ότι θέλουμε να υπολογίσουμε το εμβαδόν ενός επίπεδου σχήματος Σ. Μπορεί το σχήμα να είναι τελείως αυθαίρετο με καμπυλόγραμμα όρια, μπορεί να δίνεται γραφικά ή αναλυτικά και μπορεί να είναι είτε ενιαίο είτε να αποτελείται από πολλά τμήματα. Έστω Σ η περιοχή που είναι σχεδιασμένη στο παρακάτω σχήμα και η οποία υποθέτουμε ότι βρίσκεται ολόκληρη εντός του μοναδιαίου τετραγώνου.

Διαλέγουμε τυχαία N σημεία στο τετράγωνο και ορίζουμε ως N' τον αριθμό των σημείων που τυχαίνει να "πέφτουν" εντός του Σ. Τότε είναι γεωμετρικά προφανές ότι το εμβαδόν του Σ είναι προσεγγιστικά ίσο με το λόγο: N’ / N . Όσο μεγαλύτερο είναι το N, τόσο πιο μεγάλη είναι η ακρίβεια αυτής της εκτίμησης.



Ο αριθμός των σημείων του σχήματος 1 είναι N = 40. Απο αυτά N´ = 12 είναι μέσα στην περιοχή Σ. Ο λόγος N´/N= 12/40 = 0.30, ενώ το πραγματικό εμβαδόν του Σ είναι 0.35.

Πρακτικά, η μέθοδος Hit or Miss **δεν χρησιμοποιείται** για ευρεση του εμβαδού ενός επίπεδου σχήματος, και αυτό διότι υπάρχουν μέθοδοι - τύποι σύγκλισης που αν και είναι πιο πολύπλοκοι μαθηματικά, είναι πολύ πιο ακριβείς,

Παρόλα αυτά η μέθοδος "Επιτυχίας ή Αστοχίας" του παραδείγματος αυτού, μας επιτρέπει να εκτιμήσουμε, το ίδιο απλά, το "πολυδιάστατο όγκο" ενός σώματος σε πολλές διαστάσεις. Σε αυτήν την περίπτωση η MC είναι συχνά η μόνη αριθμητική μέθοδος που είναι χρήσιμη στην επίλυση του προβλήματος.

1.6 Ποιά είναι τα τρία κυριότερα χαρακτηριστικά των μεθόδων MC

- **Η Monte Carlo είναι πολύ απλή**. Ενα στοιχείο που χαρακτηρίζει την MC είναι η απλή δομή του υπολογιστικού αλγόριθμου. Σαν κανόνας, το πρόγραμμα πρέπει να διεκπεραιώσει μια τυχαία δοκιμή (στο προηγούμενο παράδειγμα της "Επιτυχίας ή Αστοχίας" πρέπει κανείς να ελέγξει αν ένα επιλεγμένο τυχαίο σημείο του τετραγώνου βρίσκεται επίσης μέσα στην περιοχή Σ). Αυτή η δοκιμή επαναλαμβάνεται N φορές, με κάθε δοκιμή ανεξάρτητη από τις προηγούμενες, και μετά τα αποτελέσματα των δοκιμών παίρνονται κατά μέσο όρο.

- **Η μέθοδος μπορεί να είναι εξαιρετικά αργή. Για παράδειγμα στην ολοκλήρωση, το λάθος των υπολογισμών είναι ανάλογο με $\sqrt{D/N}$, όπου D είναι κάποια σταθερά και N ο αριθμός των δοκιμών**. Γίνεται προφανές έτσι ότι για να μειώσει κανείς το λάθος κατά ένα παράγοντα 10 (δηλαδή να αποκτήσει ακόμη ένα δεκαδικό ψηφιο) απαιτείται να αυξήσει το N κατά 100 (και την ποσότητα εργασίας)

- Η μέθοδος **είναι πολύ δυνατή στα πολυδιάστατα προβλήματα**, όπου και χρησιμοποιείται κυρίως, γιατί γενικά η ακρίβεια της εξαρτάται μόνο από την πολυπλοκότητα του προβλήματος. Στην ολοκλήρωση, όπως θα δούμε αργότερα, συγκλινει με ρυθμό $O(N^{-1/2})$ που είναι ανεξάρτητο των διαστάσεων του ολοκληρώματος. Για αυτό το λόγο, η μέθοδος MC είναι η μόνη βιώσιμη σε ένα μεγάλο πεδίο προβλημάτων πολλών διαστάσεων, από την φυσική ως την οικονομία.

Λόγω των παραπάνω χαρακτηριστικών της μεθόδου, της **εύκολης χρήσης, της εφαρμοσιμότητας σε πολλά πεδία, αλλά και της αργής σύγκλισης** ένα μεγάλο ποσο υπολογιστικού χρόνου δαπανάται σε υπολογισμούς με την Monte Carlo.

Για αυτό το λόγο, ακόμα και μέτριες βελτιώσεις στην μέθοδο μπορεί να έχουν ουσιαστικό αντίκτυπο στην αποτελεσματικότητα και την εφαρμοσιμότητά της.

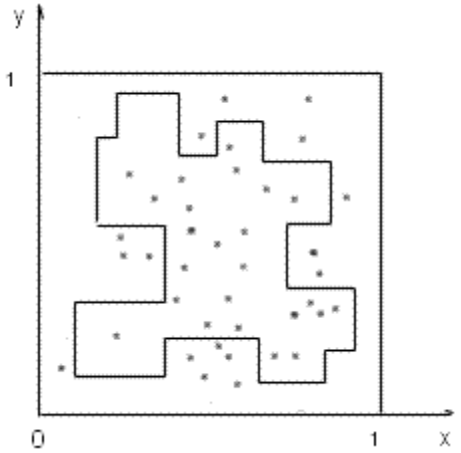
1.7 Πως προσπαθούμε να βελτιώσουμε την μέθοδο

Ο κύριος τρόπος για να βελτιώσουμε την ταχύτητα των υπολογισμών Monte Carlo είναι η μείωση της διασποράς (Variance Reduction). Οι μέθοδοι Μείωσης της Διασποράς της Monte Carlo αποσκοπούν να επιταχύνουν τον ρυθμό σύγκλισης μειώνοντας την σταθερά μπροστά από τον $O(N^{-1/2})$ χρησιμοποιώντας (ψευδο-) τυχαίες (pseydo-random) ακολουθίες αριθμών. Μια εναλλακτική προσέγγιση είναι να αλλάξουμε ακολουθίες από (ψευδο-) τυχαίους αριθμούς σε quasi-τυχαίους αριθμούς (quasi-random). Το χαρακτηριστικό αυτών είναι όπως θα δούμε πως δεν προσπαθούν να μιμηθούν την συμπεριφορά των

τυχαίων ακολουθιών, αλλά τα στοιχεία των quasi-τυχαίων ακολουθιών είναι συσχετισμένα ώστε να γίνουν ηθελημένα πιο ομοιόμορφα από τις «καθαρά» τυχαίες ακολουθίες.

1.8 Περισσότερα για το παράδειγμα Επιτυχίας ή Αστοχίας

Πώς γίνεται η επιλογή των σημείων στο παράδειγμα αυτό; Έστω ότι μια μεγενθυμένη εκδοχή του σχήματος 1 κρέμεται σε έναν τοίχο σαν στόχος. Σε κάποια απόσταση από τον τοίχο, ένας σκοπευτής πυροβολεί N φορές, στοχεύοντας στο κέντρο το τετραγώνου. Υποθέτουμε ότι η απόσταση από τον τοίχο είναι αρκετά μεγάλη (και ότι ο σκοπευτής δεν είναι ο παγκόσμιος πρωταθλητής…). Τότε, φυσικά, οι σφαίρες δεν θα πετύχουν το κέντρο, αλλά θα χτυπήσουν τον στόχο σε N τυχαία σημεία. Μπορούμε να βρούμε τότε το εμβαδόν του Σ από αυτά τα σημεία;



Το αποτέλεσμα ενός τέτοιου πειράματος φαίνεται στο παραπάνω σχήμα. Σε αυτό, έχουμε N = 40, N΄ = 24 και ο λόγος είναι N΄/N = 0.60, που είναι σχεδόν διπλάσιο από το πραγματικό εμβαδόν (0.35). Γίνεται τότε προφανές ότι εάν ο σκοπευτής ήταν πάρα πολύ ικανός, το αποτέλεσμα θα ήταν πολύ παραπλανητικό, διότι σχεδόν όλες οι σφαίρες θα χτυπούσαν τον στόχο κοντά στο κέντρο και φυσικά εντός του Σ.

Αντιλαμβάνεται, λοιπόν, κανείς ότι η μέθοδος του υπολογισμού εμβαδού είναι έγκυρη, όταν τα τυχαία σημεία είναι όχι μόνο "τυχαία" αλλά επιπλέον και "ομοιόμορφα κατανεμημένα" σε όλο το τετράγωνο.

2. Στοιχεία Στατιστικής στην Monte Carlo

2.1 Νόμος των Μεγάλων αριθμών

Έστω μια διακριτή τυχαία μεταβλητή Χ. Εάν παρατηρήσουμε την μεταβλητή Χ αρκετές φορές και πάρουμε τις τιμές $\xi_1, \xi_2, \ldots, \xi_N$, που κάθε μια θα είναι ίση με ένα από τους αριθμούς $x_1, x_2, x_3, \ldots, x_n$ που είναι οι δυνατές τιμές της Χ, τότε ο αριθμητικός μέσος όρος αυτών των τιμών θα είναι κοντά στην μέση τιμή *E(X)*, δηλαδή :

$$\frac{1}{N}(\xi_1 + \xi_2 + \ldots + \xi_N) \approx E(X)$$

και η διασπορά Var(*X*) χαρακτηρίζει την διακύμανση των τιμών αυτών γύρω από τη μέση τιμή *E(X)*. Επίσης για $N \to \infty$ θα έχουμε ότι $\lim \frac{1}{N}\sum_{j=1}^{N} \xi_j = E(X)$

Η παραπάνω εξίσωση είναι μια απλή περίπτωση του *Νόμου των μεγάλων αριθμών* και μπορεί να επεξηγηθεί ως εξής: Υποθέτουμε ότι μεταξύ των τιμών $\xi_1, \xi_2, \ldots, \xi_N$, ο αριθμός x_1 εμφανίζεται k_1 φορές, ο αριθμός x_2 εμφανίζεται k_2 φορές , … κ.ο.κ. και ο αριθμός x_n εμφανίζεται k_n φορές. Τότε:

$$\sum_{j=1}^{N} \xi_j = x_1 k_1 + x_2 k_2 + \ldots + x_n k_n$$

Έτσι:

$$\frac{1}{N}\sum_{j=1}^{N} \xi_j = x_1 \frac{k_1}{N} + x_2 \frac{k_2}{N} + \ldots + x_n \frac{k_n}{N}$$

Για μεγάλο *N*, η συχνότητα k_i / N της τιμής x_i προσεγγίζει την πιθανότητα της p_i ,έτσι ώστε $k_i / N \approx p_i$. Ετσι θα έχουμε :

$$\frac{1}{N}\sum_{j=1}^{N} \xi_j = \sum_{i-1}^{n} x_i \frac{k_i}{N} \approx \sum_{i=1}^{n} x_i p_i = E(X)$$

2.2 Η κατανομή Gauss και ο κανόνας των 3 αποκλίσεων

Μια κανονική (ή Gauss) τυχαία μεταβλητή είναι μια τυχαία μεταβλητή Χ που ορίζεται στο $\Re$ και έχει πυκνότητα πιθανότητας:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

όπου μ και σ > 0 είναι πραγματικές παράμετροι.
Αποδεικνύεται ότι η μέση τιμή και η διασπορά μιας κανονικής τυχαίας μεταβλητής είναι αντίστοιχα: $E(X) = \mu \quad \kappa\alpha\iota \quad Var(X) = \sigma^2$

Η συνάρτηση $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ λέγεται συνάρτηση σφάλματος (error function) και συνήθως χρησιμοποιείται ένας πίνακας τιμών της για να πάρουμε αυθαίρετες πιθανότητες μιας κανονικής τυχαίας μεταβλητής X . Αυτό αποδεικνύεται ως εξής:
Έστω ότι θέλουμε να βρούμε την πιθανότητα:

$$P(a < X < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_\alpha^b \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

Κάνουμε την αντικατάσταση: $x - \mu = \sigma t$ και παίρνουμε:

$$P(a < X < b) = \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} \exp\left(-\frac{t^2}{2}\right) dt$$

όπου $t_1 = (a-\mu)/\sigma \quad \kappa\alpha\iota \quad t_2 = (\beta-\mu)/\sigma$

Έτσι φτάνουμε στη σχέση:

$$P(a < X < b) = \frac{1}{2}\left(\Phi(t_2) - \Phi(t_1)\right)$$

που χρησιμοποιούμε για να πάρουμε τις ζητούμενες πιθανότητες της X. Οι πινακες τιμών της Φ περιέχουν μόνο θετικές τιμές του x, αφού ισχύει: Φ(-x) = -Φ(x).

Έστω τώρα ότι θέτουμε α=μ-3σ και b=μ+3σ. Τότε $t_1 = -3 \quad \kappa\alpha\iota \quad t_2 = 3$ και η σχέση που μας δίνεται την πιθανότητα της X γίνεται:

$$P(\mu - 3\sigma < \mathrm{X} < \mu + 3\sigma) = \Phi(3) = 0.997$$

Δηλαδή με πιθανότητα 0.997 (σχεδόν 1), μια τυχαία δοκιμή από την κανονική κατανομή θα μας δώσει μια τιμή της X που δεν θα απέχει από την μέση τιμή της περισσότερο από τρεις αποκλίσεις («κανόνας των 3 σ»).

2.3 Κεντρικό Οριακό Θεώρημα

Ας θεωρήσουμε *N* ίδιες ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, ..., X_N$, έτσι ώστε οι κατανομές πιθανοτήτων τους να συμπίπτουν. Συμπερασματικά, οι μέσες τους τιμές και οι διασπορές θα συμπίπτουν επίσης (υποθέτουμε ότι είναι πεπερασμένες). Αυτές οι τυχαίες μεταβλητές μπορούν να είναι συνεχείς ή διακριτές.

Ας ορίσουμε:

$$E(X_1) = E(X_2) = ... = E(X_N) = m$$
$$Var(X_1) = Var(X_2) = ... = Var(X_N) = b^2$$

Δηλώνουμε το άθροισμα όλων των μεταβλητών με S_N:

$$S_N = X_1 + X_2 + ... + X_N$$

Από τις ιδιότητες $E(\mathrm{X}+\mathrm{Y}) = E(\mathrm{X}) + E(\mathrm{Y})$ και $Var(\mathrm{X}+\mathrm{Y}) = Var(\mathrm{X}) + Var(\mathrm{Y})$ έχουμε ότι:

$$E(S_N) = E(\mathrm{X}_1 + \mathrm{X}_2 + ... + \mathrm{X}_N) = E(\mathrm{X}_1) + E(\mathrm{X}_2) + ... + E(\mathrm{X}_N) = Nm$$

και

$$Var(S_N) = Var(\mathrm{X}_1 + \mathrm{X}_2 + ... + \mathrm{X}_N) = Var(\mathrm{X}_1) + Var(\mathrm{X}_2) + ... + Var(\mathrm{X}_N) = Nb^2$$

Ας θεωρήσουμε τώρα μια κανονική τυχαία μεταβλητή Z_N με τις ίδιες παραμέτρους, δηλαδή $\mu = Nm$ και $\sigma = b\sqrt{N}$. Η πυκνότητα πιθανότητας της συμβολίζεται με: $p_N(x)$.

Τότε το *Κεντρικό Οριακό Θεώρημα* (Κ.Ο.Θ.) λέει ότι για κάθε διάστημα (x', x'') και για όλα τα μεγάλα *N* έχουμε:

$$P(x' < S_N < x'') \approx \int_{x'}^{x''} p_N(x)dx$$

Η φυσική σημασία του θεωρήματος αυτού είναι προφανής: *το άθροισμα S_N ενός μεγάλου αριθμού ίδιων ανεξάρτητων τυχαίων μεταβλητών είναι προσεγγιστικά μια κανονική τυχαία μεταβλητή.* Στην πραγματικότητα, αυτό το θεώρημα ισχύει και κάτω από πολύ πιο αδύναμες συνθήκες: οι μεταβλητές $X_1, X_2, ..., X_N$ να μην είναι απαραίτητα πανομοιότυπες και ανεξάρτητες, ουσιαστικά δηλαδή αυτό που απαιτείται είναι οι μεμονωμένες μεταβλητές X_j να μην παίζουν πολύ μεγάλο ρόλο στο άθροισμα.

Αυτό το θεώρημα είναι που εξηγεί γιατί οι κανονικές τυχαίες μεταβλητές συναντώνται τόσο συχνά στη φύση. **Πράγματι, όταν συναντάμε μια αθροιστική επίδραση ενός μεγάλου αριθμού μικρών τυχαίων παραγόντων, η προκύπτουσα (αθροιστική) τυχαία μεταβλητή είναι κανονική**. Για παράδειγμα, ο διασκορπισμός των όλμων πυροβολικού είναι σχεδόν πάντα κανονική τυχαία μεταβλητή, αφού εξαρτάται από τις καιρικές συνθήκες καθ'όλες τις διαφορες περιοχές της τροχιάς όπως επίσης και απο πολλούς άλλους παράγοντες.

2.4. Το γενικό σχήμα των μεθόδων Monte Carlo

Το γενικό σχήμα των μεθόδων MC βασίζεται στον Κεντρικό Οριακό Θεώρημα και στον κανόνα των 3 αποκλίσεων.
Υποθέτουμε ότι χρειάζεται να υπολογίσουμε κάποια άγνωστη ποσότητα *m*. Ας προσπαθήσουμε να βρούμε μια τυχαία μεταβλητή X με μέση τιμή *E(X)=m*. Έστω ότι η διασπορά της X είναι $Var(X) = b^2$.

Θεωρούμε *N* ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, ..., X_N$ με κατανομές πανομοιότυπες με εκείνη της X. Εάν το *N* είναι αρκετά μεγάλο, τότε έπεται από το Κεντρικό Οριακό Θεώρημα ότι η κατανομή του αθροίσματος

$$S_N = X_1 + X_2 + ... + X_N$$

θα είναι προσεγγιστικά κανονική, με παραμέτρους: $\mu = Nm$ και $\sigma = b\sqrt{N}$.
Σύμφωνα με τον κανόνα των "τριών σ" θα έχουμε:

$$P\left(Nm - 3b\sqrt{N} < S_N < Nm + 3b\sqrt{N}\right) \approx 0.997$$

Αν διαιρέσουμε την ανισότητα της παρένθεσης με *N*, παίρνουμε μια ισοδύναμη ανισότητα, της οποίας η πιθανότητα παραμένει η ίδια:

$$P\left(m - \frac{3b}{\sqrt{N}} < \frac{S_N}{N} < m + \frac{3b}{\sqrt{N}}\right) \approx 0.997$$

Ξαναγράφουμε την τελευταία έκφραση σε κάπως διαφορετική μορφή:

$$P\left(\left|\frac{1}{N}\sum_{j=1}^{N} X_j - m\right| < \frac{3b}{\sqrt{N}}\right) \approx 0.997$$

Αυτή είναι μια πολύ σημαντική σχέση για την μέθοδο MC, που μας δίνει και την μέθοδο υπολογισμού της *m* αλλά και μια προσέγγιση του λάθους.

Πράγματι, αρκεί να πάρουμε *N* τιμές της τυχαίας μεταβλητής X - η επιλογή μιας τιμής από κάθε μια από τις $X_1, X_2, ..., X_N$ είναι ισοδύναμη με την επιλογή *N* τιμών της X, αφού όλες αυτές οι τυχαίες μεταβλητές έχουν ίδιες κατανομές και να πάρουμε τη δειγματική μέση τιμή τους. Για αυτό και στην βιβλιογραφία, αυτό το σχήμα λέγεται και Sample Mean Monte Carlo.

Από την παραπάνω εξίσωση, γίνεται φανερό ότι ο αριθμητικός μέσος αυτών των τιμών θα είναι προσεγγιστικά ίσος με τη ζητούμενη ποσότητα *m*. Με κάθε πιθανότητα, το σφάλμα της προσέγγισης δεν ξεπερνά το $3b/\sqrt{N}$ και προσεγγίζει το μηδεν καθώς το *N* αυξάνει.

3. Ολοκλήρωση Monte Carlo

3.1 Η μέθοδος του υπολογισμού

Θεωρούμε μια συνάρτηση $f(x)$ ορισμένη στο διάστημα *α < x < b.* Θέλουμε να υπολογίσουμε προσεγγιστικά το ολοκλήρωμα:

$$\mathbf{I} = \int_a^b f(x)dx$$

Ξεκινάμε επιλέγοντας μια αυθαίρετη πυκνότητα κατανομής *p(x)* ορισμένη στο διάστημα *(α, b)* - με άλλα λόγια μια αυθαίρετη συνάρτηση που ικανοποιεί τις εξισώσεις

1. $p(x) > 0$
2. $\int_a^b p(x)dx = 1$

Θεωρούμε μια τυχαία μεταβλητή X με πυκνότητα *p(x)* στο διάστημα *(α, b)*, και επιπλέον μια τυχαία μεταβλητή

$$Y = g(X) = f(X)/p(X)$$

Από την ιδιότητα: $E(g(X)) = \int_a^b g(x)p(x)dx$ θα έχουμε: $E(Y) = \int_a^b \frac{f(x)}{p(x)} p(x)dx = \mathbf{I}$

Τώρα, θεωρούμε *N* ανεξάρτητες, τυχαίες μεταβλητές $Y_1, Y_2, ..., Y_N$ με την ίδια κατανομή με την Y και εφαρμόζουμε το Κεντρικό Οριακό Θεώρημα για το άθροισμά τους, το οποίο θα είναι μια τυχαία μεταβλητή προσεγγιστικά κανονική με παραμέτρους $a = N \cdot \mathbf{I}$, $\sigma = \sqrt{N \cdot Var(Y)}$ Σε αυτήν την περίπτωση θα έχουμε από τον κανόνα των 3σ:

$$\mathbf{P}(a - 3\sigma < S_N < a + 3\sigma) \approx 0.997$$

$$\mathbf{P}\left(N \cdot \mathbf{I} - 3\sqrt{N \cdot Var(Y)} < S_N < N \cdot \mathbf{I} + 3\sqrt{N \cdot Var(Y)}\right) \approx 0.997$$

$$\mathbf{P}\left(\mathbf{I} - \frac{3\sqrt{N \cdot Var(Y)}}{N} < \frac{S_N}{N} < \mathbf{I} + \frac{3\sqrt{N \cdot Var(Y)}}{N}\right) \approx 0.997$$

$$\mathbf{P}\left(\left|\frac{1}{N}\sum_{j=1}^{N} Y_j - \mathbf{I}\right| < 3 \cdot \sqrt{\frac{Var(Y)}{N}}\right) \approx 0.997 \qquad (\mathbf{II})$$

Από την τελευταία σχέση, έχουμε ότι εάν διαλέξουμε *N* τιμές $\xi_1, \xi_2, ..., \xi_N$, τότε για *N* αρκετά μεγάλο θα έχουμε ότι:

$$\frac{1}{N}\sum_{j=1}^{N} \frac{f(\xi_j)}{p(\xi_j)} \approx I \qquad (\mathbf{III})$$

Η σχέση (**II**) επίσης μας δίνει και ένα πιθανοθεωρητικό άνω φράγμα του σφάλματος της προσέγγισης της (**III**), το οποίο θα είναι $3 \cdot \left(\frac{Var(Y)}{N} \right)^{\frac{1}{2}}$. Για αυτό και λέμε ότι το σφάλμα της Monte Carlo ολοκλήρωσης είναι της τάξης $O\left(N^{\frac{1}{2}} \right)$.

3.2 Το σχήμα της ολοκλήρωσης όταν πάρουμε ομοιόμορφη τ.μ.

Θεωρούμε το μονοδιάστατο ολοκλήρωμα $I = \int_0^1 f(x)dx$.

Σύμφωνα λοιπόν με το σχήμα των μεθόδων Monte Carlo, εάν θέλουμε να υπολογίσουμε το ολοκλήρωμα της συνάρτησης $f(x)$, αρκεί να βρούμε μια τυχαία μεταβλητή ***X*** της οποίας η μέση τιμή να είναι ίση με ολοκλήρωμα της $f(x)$.

Έστω ***X*** να είναι μια ομοιόμορφα κατανεμημένη τυχαία μεταβλητή στο [0,1], δηλαδή η πυκνότητα της θα είναι ίση με $\phi(x) = 1$. Η μέση ή αναμενόμενη τιμή τότε της $f(x)$ στο διάστημα [0,1] είναι:

$$E_X[f(x)] = \int_0^1 f(x)\phi(x)dx$$

$$= \int_0^1 f(x)dx$$

όπου η $E_X[]$ είναι η μέση τιμή σε σχέση με την τυχαία μεταβλητή ***X***. Έστω x_i μια τυχαία δοκιμή από την κατανομή ***X ~ U[0,1]***. Γνωρίζουμε από τον νόμο των μεγάλων αριθμών ότι :

$$E_X[f(x)] = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

Άρα μπορούμε να προσεγγίσουμε το ολοκλήρωμα $\int_0^1 f(x)dx$ χρησιμοποιώντας τυχαίες δοκιμές-αριθμούς από την ***X*** σύμφωνα με τον τύπο:

$$\hat{I}_n = \frac{1}{N} \sum_{i=1}^{N} f(x_i) \approx \int_0^1 f(x)dx$$

Σύμφωνα με τον ισχυρό Νόμο των μεγάλων αριθμών αυτή η προσέγγιση συγκλίνει με πιθανότητα 1. Δηλαδή

$$\lim_{N \to \infty} \hat{I}_N = I$$

Αφού η προσέγγιση βασίζεται σε μια τυχαία μεταβλητή, το προσεγγιστικό ολοκλήρωμα είναι και αυτό μια τυχαία μεταβλητή με διασπορά ίση με:

$$\begin{aligned} V(\hat{I}_N) &= \frac{1}{N} V(f(x)) \\ &= \frac{1}{N} E\left[\left(f(x) - E[f(x)]\right)^2\right] \\ &= \frac{1}{N} E\left[\left(f(x) - I\right)^2\right] \end{aligned}$$

Ένας αμερόληπτος εκτιμητής της $V(f(x))$ είναι ως γνωστόν:

$$\hat{V}(f(x)) = \frac{1}{N-1} \sum_{i=1}^{N} \left(f(x_i) - \hat{I}_N\right)^2$$

Έτσι η εκτιμώμενη διασπορά του εκτιμητή του ολοκληρώματος (της προσέγγισης) είναι:

$$\hat{V}(\hat{I}_N) \approx \frac{1}{N} \frac{1}{N-1} \sum_{i=1}^{N} \left(f(x_i) - \hat{I}_N\right)^2$$

<u>Παρατήρηση:</u> Πρέπει να τονίσουμε ότι κάθε προσέγγιση ολοκλήρωσης Monte Carlo θα πρέπει να συνοδεύεται με αναφορά για την διασπορά. Αλλιώς θα είναι σαν να αναφέρουμε μια εκτίμηση χωρίς το τυπικό σφάλμα, δηλαδή πάνω κάτω άσκοπο. Πρέπει λοιπόν να γνωρίζουμε την τάξη του προσεγγιστικού σφάλματος.

3.3 Το σφάλμα της ολοκλήρωσης Monte Carlo

Ορίζουμε το σφάλμα της Monte Carlo σαν:

$$\varepsilon_{\mathrm{N}}[f] = I - \hat{I}_N ,$$

δηλαδή

$$\varepsilon_N[f] = \int_{I^d} f(x)dx - \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

και το το μέσο τετραγωνικό σφάλμα (Root Mean Square Error (RMSE)) της MC ολοκλήρωσης που χρησιμοποιεί (ψευδο-)τυχαίες ακολουθίες: $E\left[\varepsilon[f]^2\right]^{1/2}$

Το Κεντρικό Οριακό Θεώρημα περιγράφει το μέγεθος και τις στατιστικές ιδιότητες των σφαλμάτων της ολοκλήρωσης Monte Carlo.

<u>Θεώρημα:</u> Για N μεγάλο, έχουμε

$$\varepsilon_{\mathrm{N}}[f] \approx \sigma\, \mathrm{N}^{-1/2} \nu$$

όπου ν είναι η τυπική (N(0,1)) τυχαία μεταβλητή και η σταθερά σ είναι η τυπική απόκλιση, δηλαδή η τετραγωνική ρίζα της διασποράς της f :

$$\sigma = \left(\int_{x \in D} \left(f(x) - I \right)^2 dx \right)^{1/2}$$

Το θεώρημα αυτό μας λέει ότι το σφάλμα της Monte Carlo είναι της τάξης του $O(N^{-1/2})$ με μια σταθερά μπροστά που είναι ακριβώς η διασπορά της συνάρτησης f. Επιπλέον η κατανομή του σφάλματος είναι προσεγγιστικά κανονική.

Αφού το φράγμα του σφάλματος από το Κ.Ο.Θ. είναι πιθανοθεωρητικό, η ακρίβεια της μεθόδου μπορεί να διασφαλιστεί με κάποιο επίπεδο εμπιστοσύνης c. Για να είμαστε σίγουροι ότι το σφάλμα είναι μεγέθους το πολύ ε με κάποιο επίπεδο εμπιστοσύνης c, απαιτείται ο αριθμός των σημείων N που χρησιμοποιούμε να είναι:

$$N = \varepsilon^{-2} \sigma^2 s(c)$$

όπου s είναι η συνάρτηση εμπιστοσύνης (confidence function) για μια κανονική τυχαία μεταβλητή. Δηλαδή:

$$c = \int_{-s(c)}^{s(c)} e^{-c^2/2} dx / \sqrt{2\pi} = erf\left(s(c) / \sqrt{2} \right)$$

Στην πράξη είναι δύσκολο να βρούμε την ακριβή τιμή της διασποράς, άρα και της απόκλισης, για αυτό οι παραπάνω τύποι δεν εφαρμόζονται άμεσα. Γι ‘ αυτό χρησιμοποιούμε την εμπειρική διασπορά και σφάλμα. Επαναλαμβάνουμε το πείραμα Monte Carlo M φορές χρησιμποιώντας ανεξάρτητα σημεία x_i για $1 \le i \le MN$. Για κάθε επανάληψη παίρνουμε τιμές $I_N{}^{(j)}$ για $1 \le j \le M$. Τότε το εμπειρικό μέσο τετραγωνικό σφάλμα είναι :

$$\bar{\varepsilon}_{\mathrm{N}} = \left(\frac{1}{\mathrm{M}} \sum_{j=1}^{M} \left(I_N{}^{(j)} - \bar{I}_N \right)^2 \right)^{1/2}$$

όπου έχουμε:

$$\bar{I}_N = \frac{1}{M} \sum_{j=1}^{M} I_N{}^{(j)}$$

τότε η εμπειρική διασπορά δίνεται από τη σχέση:

$$\bar{\sigma} = \mathrm{N}^{1/2} \bar{\varepsilon}_{\mathrm{N}}$$

Αυτήν την τιμή μπορούμε να χρησιμοποιήσουμε στην σχέση $N = \varepsilon^{-2}\sigma^2 s(c)$ για να βρούμε το πλήθος των σημείων που απαιτούνται για μια δεδομένη ακρίβεια ε και ένα επίπεδο εμπιστοσύνης c.

3.4 Επέκταση στην πολυδιάστατη περίπτωση

Η επέκταση στην γενική πολυδιάστατη περίπτωση είναι αρκετά απλή. Έστω $D \subset \Re^n$ να είναι ένα ν-διάστατο πεδίο πάνω στο οποίο θέλουμε να υπολογίσουμε ένα ολοκλήρωμα $\int_{x \in D} f(\mathbf{x})d\mathbf{x}$ όπου $f(x): \Re^n \to \Re$. Έστω $X \in D$ να είναι μια τυχαία μεταβλητή πάνω στο πεδίο *D* με ομοιόμορφη κατανομή $\phi(\mathbf{x}) = \frac{1}{vol(D)}$ έτσι ώστε να ισχύει: $\int_{x \in D} \phi(\mathbf{x})d\mathbf{x} = 1$. Τότε το ολοκλήρωμα είναι μια αναμενόμενη τιμή:

$$\int_{\mathbf{x} \in D} f(\mathbf{x})d\mathbf{x} = \int_{\mathbf{x} \in D} \frac{f(\mathbf{x})}{\phi(\mathbf{x})}\phi(\mathbf{x})d\mathbf{x} =$$

$$= E\left(\frac{f(\mathbf{x})}{\phi(\mathbf{x})}\right) = vol(D)\ E(f(\mathbf{x})) \approx$$

$$\approx \frac{vol(D)}{n}\sum_{i=1}^{n} f(\mathbf{x}_i)$$

όπου x_i να είναι μια τυχαία δοκιμή από εκείνη την κατανομή, δηλαδή ένα διάνυσμα τυχαίων τιμών.

Τα κυριότερα προβλήματα στην ολοκλήρωση Monte Carlo είναι:

Τυχαίες Δοκιμές: Πρέπει να δοκιμάζουμε με τυχαίους αριθμούς από την "σωστή" κατανομή.

Μείωση Διασποράς: Για να ελαχιστοποιήσουμε το σφάλμα της προσέγγισης (τη διασπορά του εκτιμητή).

3.5 Σύγκριση με μεθόδους που βασίζονται σε πλέγμα (grid based methods)

Πως μπορεί μια μέθοδος που βασίζεται σε τυχαίους αριθμούς να είναι καλύτερη από τις κλασσικές μεθόδους που βασίζονται σε πλέγματα; Ας συγκρίνουμε τον ρυθμό σύγκλισης της Monte Carlo με αυτόν μιας αριθμητικής μεθόδου ολοκλήρωσης, π.χ. της μεθόδου Simpson. Ο ρυθμός σύγκλισης μιας μεθόδου του είδους Newton-Cotes είναι $O\left(N^{-k/d}\right)$, όπου k η τάξη της μεθόδου και d ο

αριθμός των διαστάσεων. Ο ρυθμός σύγκλισης όμως της μεθόδου Monte Carlo είναι όπως είδαμε $O\left(N^{-1/2}\right)$ και είναι ανεξάρτητος των διαστάσεων. Έτσι, μπορούμε να πούμε ότι η Monte Carlo υπερισχύει των μεθόδων Newton-Cotes εάν

$$k/d < 1/2$$

Μια πιο ρεαλιστική εξήγηση είναι ότι είναι πρακτικά αδύνατο να δημιουργήσουμε ένα πλέγμα σε πολλές διαστάσεις. Το απλούστερο πλέγμα για έναν κύβο d διαστάσεων απαιτεί 2^d σημεία. Για 20 διαστάσεις, που δεν είναι ιδιαίτερα μεγάλος αριθμός για εφαρμογές, απαιτούνται περισσότερο από ένα εκατομμύριο σημεία. Επιπλέον είναι αδύνατο να εκλεπτύνουμε ένα πλέγμα σε πολλές διαστάσεις αφού μια τέτοια εκλέπτυνση απαιτεί αύξηση των σημείων κατά ένα παράγοντα του 2^d. Σε αντίθεση με αυτά, η ακρίβεια της Monte Carlo δεν εξαρτάται από τις διαστάσεις και κάθε επιπλέον σημείο που προστίθεται στον υπολογισμό προσδίνει μια αυξητική βελτίωση στην ακρίβειά της.

4. Παραγωγή τυχαίων αριθμών

Όπως είδαμε, στα προηγούμενα θεωρήσαμε κυρίως για απλούστευση ότι χρησιμοποιούμε τυχαίες μεταβλητές απο την ομοιόμορφη κατανομή στην ολοκλήρωση Monte Carlo. Δεν ισχύει όμως πάντα αυτό. Μπορεί να χρειάζεται να χρησιμοποιήσουμε τυχαίους αριθμούς από μια άλλη κατανομή. Οι συνήθεις γεννήτριες τυχαίων αριθμών στους υπολογιστές μας δίνουν τυχαίες μεταβλητές-αριθμούς από την ομοιόμορφη κατανομή. Για να πάρουμε τυχαίους αριθμούς από μη-ομοιόμορφες κατανομές κάνουμε μετασχηματισμό των ομοιόμορφων τυχαίων μεταβλητών. Παρακάτω παρουσιάζουμε μερικές μεθόδους για την παραγωγή μη – ομοιόμορφων τυχαιών μεταβλητών.

4.1 Μέθοδος Μετασχηματισμού

Έστω ότι θέλουμε να κάνουμε μια τυχαία δοκιμή από μια τυχαία μεταβλητή *X* που έχει συνάρτηση κατανομής $F(x)$. Ξέρουμε ότι η $F(X)$ είναι ομοιόμορφα κατανεμημένη στο [0,1], αφού $P(F(X) < u) = P(X < F^{-1}(u)) = F(F^{-1}(u)) = u$. Αυτό συνεπάγεται ότι αν η U είναι ομοιόμορφα κατανεμημένη στο [0,1], τότε η $F^{-1}(U)$ είναι κατανεμημένη σαν την X. Έτσι αρκεί να βρούμε την αντίστροφη της συνάρτησης κατανομής $F()$ και να εφαρμόσουμε αυτήν σε μια ομοιόμορφα κατανεμημένη τυχαία δοκιμή *u* για να πάρουμε μια τυχαία δοκιμή-αριθμό από την ζητούμενη κατανομή. Μπορεί, όμως, να υπάρχουν ουσιαστικά προβλήματα στην

εύρεση ή την προσέγγιση της αντίστροφης συνάρτησης της $F()$.

Παράδειγμα 4.1.1 Πως μπορούμε να πάρουμε μια τυχαία δοκιμή από την εκθετική κατανομή

$$F(x) = \lambda \cdot e^{-\lambda x}$$

Βρίσκουμε την αντίστροφή της συνάρτηση:

$$F^{-1}(u) = -\frac{1}{\lambda}\ln\left(\frac{u}{\lambda}\right)$$

Κάνοντας μια τυχαία δοκιμή u_i από την *U[0,1]* και θέτοντας $x_i = F^{-1}(u_i)$ παίρνουμε μια εκθετικά κατανεμημένη τυχαία δοκιμή x_i.

Παράδειγμα 4.1.2 Πως μπορούμε να πάρουμε τυχαίους αριθμούς από την κανονική κατανομή; Μια standard κανονική τυχαία μεταβλητή έχει πυκνότητα:

$$p(x) = \frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$$

και αθροιστική συνάρτηση κατανομής:

$$F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-t^2/2)dt = \frac{1}{2} + \frac{1}{2}erf\left(x/\sqrt{2}\right)$$

όπου erf είναι η συνάρτηση σφάλματος που ορίζεται από :

$$erf(z) = \frac{2}{\sqrt{2\pi}}\int_{0}^{z}\exp(-t^2)dt$$

Ένας τρόπος τότε για να πάρουμε μια δοκιμή x από την κανονική κατανομή είναι να εφαρμόσουμε την αντίστροφη της αθροιστικής συνάρτησης κατανομής *F* σε μια ομοιόμορφη τυχαία μεταβλητή y.
Θα έχουμε

$$y = F(x) = \frac{1}{2} + \frac{1}{2}erf\left(x/\sqrt{2}\right)$$

οπότε:

$$x = \sqrt{2}\,erf^{-1}(2y-1)$$

Για κάποιες κοινά χρησιμοποιούμενες κατανομές, όπως η κανονική κατανομή, υπάρχουν περισσότερο αποτελεσματικοί αλγόριθμοι για την παραγωγή τυχαίων δοκιμών-αριθμών. Για την κανονική περίπτωση μπορεί κανείς να χρησιμοποιήσει την μέθοδο Box-Muller.

Μέθοδος Box - Muller

Η μέθοδος Box-Muller παρέχει ένα άμεσο τρόπο για να παράγουμε κανονικές τυχαίες μεταβλητές χωρίς να αντιστρέφουμε την συνάρτηση σφάλματος.

Ξεκινώντας από δύο ομοιόμορφες μεταβλητές y_1, y_2, παίρνουμε δύο κανονικές τυχαίες μεταβλητές από τους τύπους:

$$x_1 = \sqrt{-2\log(y_1)}\cos(2\pi\, y_2)$$

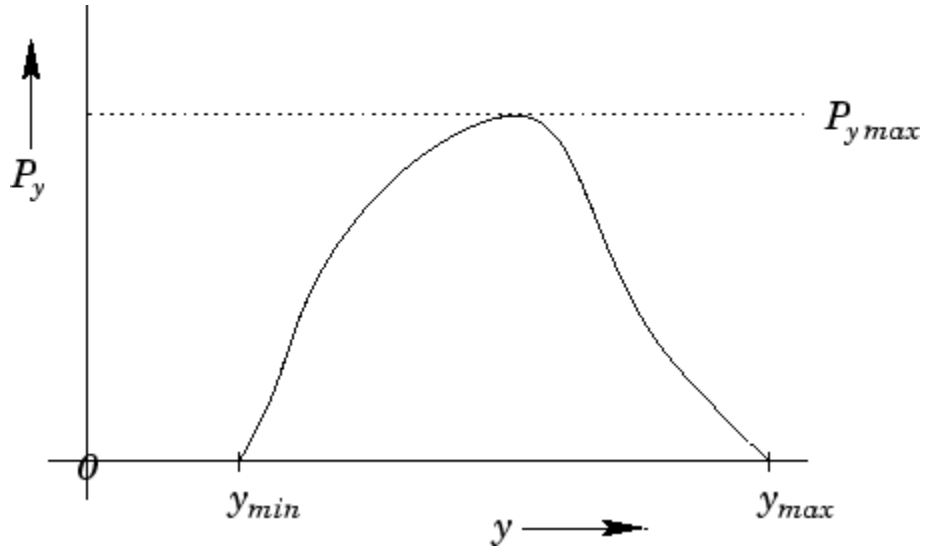
$$x_2 = \sqrt{-2\log(y_1)}\sin(2\pi\, y_2)$$

Τα μειονεκτήματα της μεθόδου Box-Muller είναι οτι απαιτεί τον υπολογισμό υπερβατικών συναρτήσεων και ότι παράγει δύο τυχαίες μεταβλητές, ανεξάρτητα αν χρειαζόμαστε μόνο μία.

4.2 Μέθοδος Αποδοχής - Απόρριψης

Έστω ότι θέλουμε να πάρουμε μια τυχαία μεταβλητή y κατανεμημένη με πυκνότητα $p_y(y)$ στο πεδίο από y_1 $\varepsilon\acute{\omega}\varsigma$ y_2. Έστω p_{MAX} η μέγιστη τιμή της πυκνότητας $p_y(y)$ στο πεδίο αυτό όπως στο σχήμα. Η μέθοδος έχει ως εξής:

- Παίρνουμε μια τυχαία τιμή y από το y_1 $\varepsilon\acute{\omega}\varsigma$ y_2
- Υπολογίζουμε την $p_y(y)$ για αυτήν την τιμή y
- Παίρνουμε ένα τυχαίο αριθμό x που είναι ομοιόμορφα κατανεμημένος στο πεδίο $[0, p_{MAX}]$
- Εάν $p_y(y) < x$ απορρίπτουμε τον αριθμό y, αλλιώς τον αποδεχόμαστε.
- Ακολουθούμε αυτή τη διαδικασία και παίρνουμε την τυχαια μεταβλητή y που είναι κατανεμημένη σύμφωνα με πυκνότητα $p_y(y)$



Παράδειγμα:
Θεωρούμε την κανονική κατανομή με πυκνότητα:

$$p_y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-(y-\mu)^2 / 2\sigma^2\right)$$

όπου μ είναι η μέση τιμή και σ η τυπική απόκλιση. Θέτουμε

$$y_{\min} = \mu - 4\sigma \quad \kappa\alpha\iota \quad y_{\max} = \mu + 4\sigma$$

αφού είναι πολύ μικρή η πιθανότητα να πάρουμε τιμές του y που να απέχουν περισσότερο από τέσσερις αποκλίσεις από την μέση της τιμή. Τότε η μέγιστη τιμή θα είναι:

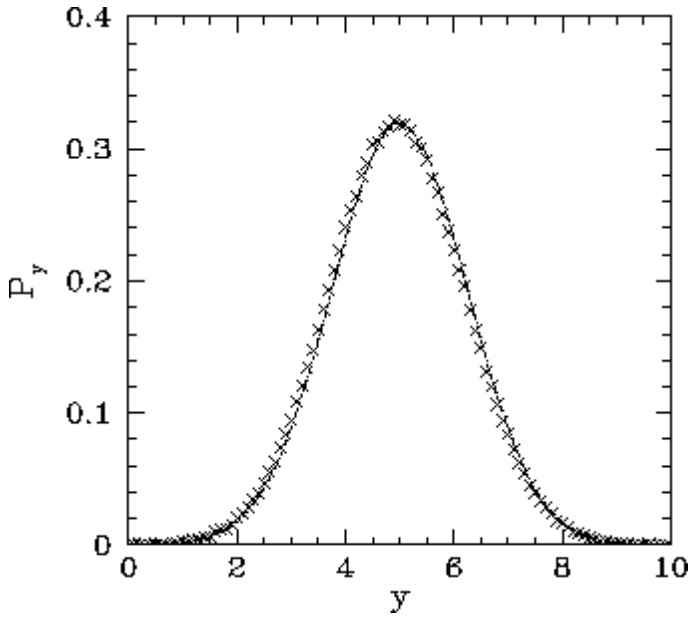
$$p_{MAX} = \frac{1}{\sigma\sqrt{2\pi}}$$

με το μέγιστο να προκύπτει όταν y=μ.

Η παρακάτω συνάρτηση της C++ επιστρέφει μια τυχαία μεταβλητή από την κανονική κατανομή με μέση τιμή mean και απόκλιση sigma:

```
  // Function to return random variable distributed
  // according to Gaussian distribution with mean mean
  // and standard deviation sigma.
 double gaussian (double mean, double sigma)
 {
 double ymin = mean - 4. * sigma;
 double ymax = mean + 4. * sigma;
 double Pymax = 1. / sqrt (2. * M_PI) / sigma;
   // Calculate random value uniformly distributed
   // in range ymin to ymax
double y = ymin + (ymax - ymin) * double (random ()) / double (RANDMAX);
   // Calculate Py
 double Py = exp (- (y - mean) * (y - mean) / 2. / sigma / sigma) / sqrt (2. * M_PI) /
sigma;
   // Calculate random value uniformly distributed in range 0 to Pymax
 double x = Pymax * double (random ()) / double (RANDMAX);
   // If x > Py reject value and recalculate
 if (x > Py) return gaussian (mean, sigma);
else return y;
}
```

Το παρακάτω σχήμα μας δείχνει την απόδοση της παραπάνω συνάρτησης για ένα εκατομμύριο σημεία με mean=5. και sigma = 1.25. Οι τιμές έχουν ομαδοποιηθεί σε 100 ομάδες με εύρος 0.1. Το σχήμα δείχνει τον αριθμό των σημείων που έπεσαν σε κάθε ομάδα διαιρεμένο κατά ένα κατάλληλο παράγοντα κανονικοποίησης. Επίσης φαίνεται η αντίστοιχη Gaussιανή καμπύλη για σύγκριση. Φαίνεται καθαρά ότι η συνάρτηση αυτή πετυχαίνει να μας δώσει μια τυχαία κανονική μεταβλητή.



Η μέθοδος έχει τα εξής πλεονεκτήματα:
1) Δεν απαιτεί τον την αντιστροφή συναρτήσεων κατανομής.
2) Δουλεύει ακόμα και αν η πυκνότητα p(x) δεν έχει κανονικοποιηθεί για να έχει ολοκλήρωμα 1.

Το μόνο της μειονέκτημα είναι ότι μπορεί να είναι αναποτελεσματική απαιτώντας πολλές επαναλήψεις πριν έχουμε μια αποδοχή.

Μια επέκταση της μεθόδου αποδοχής – απόρριψης είναι ο αλγόριθμος Metropolis.

5. Μείωση Διασποράς

Όπως είδαμε στα προηγούμενα το σφάλμα ε και ο αριθμός N των σημείων στην ολοκλήρωση Monte Carlo συσχετίζονται με τις σχέσεις:

$$\varepsilon = O(\sigma N^{-1/2})$$

$$N = O(\sigma / \varepsilon)^2$$

Ο υπολογιστικός χρόνος που απαιτείται για την MC είναι ανάλογος του αριθμού των δειγμάτων N και είναι της τάξης του $O(\sigma / \varepsilon)^2$, που μας δείχνει ότι ο

υπολογιστικός χρόνος αυξάνεται ραγδαία καθώς η απαιτούμενη ακρίβεια μεγαλώνει. Υπάρχουν δύο τρόποι για να επιταχύνουμε, να μειώσουμε το σφάλμα, της MC. Ο πρώτος είναι να προσπαθήσουμε να μειώσουμε τη διασπορά, στον οποίο το ολοκλήρωμα μετασχηματίζεται έτσι ώστε να μειωθεί η σταθερά σ. Ο δεύτερος τρόπος, που θα εξετάσουμε παρακάτω, είναι να αντικαταστήσουμε τις τυχαίες τιμές που χρησιμοποιούμε με μια εναλλακτική ακολουθία –quasi-τυχαίοι αριθμοί- που βελτιώνει τον εκθέτη ½.
Μια παρατήρηση που πρέπει να κάνουμε ότι οι μέθοδοι που ακολουθούν απαιτούν επιπλέον υπολογιστικό χρόνο, που όμως εξισορροπείται από το κέρδος που έχουμε με το μειωμένο αριθμό σημείων N. Στα περισσότερα παραδείγματα, το κέρδος λόγω της μείωσης της διασποράς είναι αρκετά σημαντικό.

5.1 Αντίθετες μεταβλητές (Antithetic Variables)

Οι αντίθετες μεταβλητές είναι μια από τις απλούστερες και πιο ευρέως χρησιμοποιούμενες μεθόδους μείωσης της διασποράς. Συνήθως οι υπολογισμοί MC χρησιμοποιούν σημεία που είναι ανεξάρτητα μεταξύ τους. Η μέθοδος των αντίθετων μεταβλητών εσκεμμένα χρησιμοποιεί συσχετισμένα σημεία, εκμεταλλευόμενη το γεγονός ότι μια τέτοια συσχέτιση μπορεί να είναι αρνητική. Μαθηματικά, βασιζόμαστε στο γεγονός ότι:

$$Var(f_1 + f_2) = Var(f_1) + Var(f_2) + 2\,coVar(f_1, f_2)$$

Εάν μπορούσαμε να διαλέξουμε σημεία τέτοια ώστε $CoVar(f_1, f_2) < 0$ τότε θα είχαμε μια ουσιαστική μείωση στην διασπορά. Για κάθε (πολυδιάστατη) δειγματική τιμή-σημείο **x**, χρησιμοποιούμε επίσης την τιμή **–x**. Έτσι η μέθοδος Monte Carlo γίνεται:

$$\hat{I}_N = \frac{1}{2N}\sum_{i=1}^{N}\left(f(\mathbf{x}_i) + f(-\mathbf{x}_i)\right)$$

Παράδειγμα:
Το πιο τετριμμένο παράδειγμα αυτής της μεθόδου είναι η ολοκλήρωση MC της συνάρτησης $f(x) = x$ στο διάστημα [0,1] υπολογίζοντας τις τιμές της στα σημεία x_1 $\kappa\alpha\iota$ $1 - x_i$. Αφού η $f(x)$ είναι μονότονη στο πεδίο ολοκλήρωσης [0,1], τότε η $CoVar(f(x), f(1-x))$ είναι αρνητική. Έτσι η ακόλουθη προσέγγιση:

$$\hat{I}_n^{\,a} = \frac{1}{n/2}\sum_{i=1}^{n/2} f(x_i) + \frac{1}{n/2}\sum_{i=1}^{n/2} f(1-x_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n/2}\left(f(x_i) + f(1-x_i)\right)$$

έχει διασπορά

$$V(\hat{I}_n^{\,a}) = \frac{1}{n}V\left(f(x)\right) + CoVar\left(f(x), f(1-x)\right)$$

που είναι μικρότερη από τον πρώτο εκτιμητή εάν $CoVar(f(x), f(1-x)) < 0$.

5.2 Μεταβλητές Ελέγχου (Control Variables)

Τώρα έστω ότι έχουμε μια εύκολα ολοκληρώσιμη συνάρτηση $g(x)$ που είναι παρόμοια με την προς ολοκληρωση συνάρτηση $f(x)$ και έχει το ίδιο πεδίο ολοκλήρωσης με αυτήν D. Τότε θα μπορούσαμε να χρησιμοποιήσουμε αυτό το γεγονός για να αυξήσουμε την διασπορά της ολοκλήρωσης Monte Carlo. Χρησιμοποιούμε την ταυτότητα:

$$I = \int_{x\in D} f(x)dx = \int_{x\in D} g(x)dx + \int_{x\in D} \big(f(x) - g(x)\big)dx \qquad (23)$$

Το πρώτο ολοκλήρωμα μπορεί να υπολογιστεί χωρίς προσέγγιση, οπότε χρειάζεται να χρησιμοποιήσουμε ολοκλήρωση Monte Carlo μόνο στο δεύτερο ολοκλήρωμα.Η διασπορά της προσέγγισης αυτής είναι τότε:

$$V(\hat{I}_n^{\,c}) = V(f) + V(g) - 2Cov(f,g)$$

Έτσι, εάν οι συναρτήσεις *g και f* είναι πολύ παρόμοιες, συνεπάγεται ότι η συνδιασπορά $Cov(f,g)$ είναι μεγάλη, και τότε $V\left(\hat{I}_n^{\,c}\right) < V\left(\hat{I}_n\right)$.

Η ιδέα των μεταβλητών ελέγχου είναι να χρησιμοποιήσουμε μια συνάρτηση g, που είναι παρόμοια με την προς ολοκλήρωση συνάρτηση f, και για την οποία το ολοκλήρωμα $I[g] = \int g(x)dx$ είναι γνωστό. Τότε το ολοκλήρωμα της f μπορεί να γραφτεί ως

$$I[f] = \int (f(x) - g(x))dx + \int g(x)dx$$

Και έτσι η μέθοδος Monte Carlo γίνεται :

$$\hat{I}_N = \frac{1}{N}\sum_{i=1}^{N}\big(f(x_i) - g(x_i)\big) + I[g]$$

της οποίας το σφάλμα $\varepsilon_{\mathrm{N}}[f] = I[f] - \hat{I}_N$ είναι μεγέθους:

$$\varepsilon_{\mathrm{N}}[f] \approx \sigma_{f-g} N^{-1/2}$$

όπου $$\sigma_{f-g}{}^2 = \int \left(\bar{f}(x) - \bar{g}(x)\right)^2 dx$$

χρησιμοποιώντας την σήμανση: $\bar{f}(x) = f(x) - I[f], \quad \bar{g}(x) = g(x) - I[f]$.
Καταλήγουμε έτσι ότι η μέθοδος αυτή είναι αποτελεσματική αν:

$$\sigma_{f-g} << \sigma_f$$

5.3 Matching Moments Μέθοδος

Το σφάλμα της ολοκλήρωσης κατά Monte Carlo οφείλεται εν μέρει στο

στατιστικό δειγματικό σφάλμα, δηλαδή στις διαφορές μεταξύ της απαιτούμενης συνάρτησης πυκνότητας p(x) και της εμπειρικής συνάρτησης πυκνότητας των δειγματισμένων σημείων $\{x_i\}_{i=1}^{N}$. Τμήμα αυτής της διαφοράς μπορεί να ιδωθεί απευθείας με την σύγκριση των στιγμών (moments) των δύο κατανομών. Ορίζουμε τις πρώτες και δεύτερες στιγμές m_1 και m_2 της p ως:

$$m_1 = \int xp(x)dx, \qquad m_2 = \int x^2 p(x)dx$$

Οι πρώτες και οι δεύτερες στιγμές μ_1 και μ_2 του δείγματος $\{x_i\}_{i=1}^{N}$ είναι:

$$\mu_1 = N^{-1}\sum_{i=1}^{N} x_i, \quad \mu_2 = N^{-1}\sum_{i=1}^{N} x_i^{\,2}$$

Το σφάλμα των στιγμών οφείλεται στην ανισότητα αυτών των στιγμών, δηλαδή:

$$\mu_1 \neq m_1, \qquad \mu_2 \neq m_2$$

Μια μερική διόρθωση του στατιστικού δειγματικού σφάλματος είναι να κάνουμε τις στιγμές να ταιριάζουν ακριβώς. Αυτό μπορεί να γίνει με έναν απλό μετασχηματισμό των δειγμάτων-σημείων. Για να ταιριάξουμε την πρώτη στιγμή του δείγματος με αυτήν της p, αντικαθιστούμε τα x_i με:

$$y_i = (x_i - \mu_1) + m_1$$

Αυτό ικανοποιεί την

$$N^{-1}\sum_{i=1}^{N} y_i = m_1$$

έτσι ώστε η πρώτη στιγμή να είναι ακριβώς σωστή. Για να ταιριάξουμε τις δύο πρώτες στιγμές, αντικαθιστούμε τα x_i με:

$$y_i = (x_i - \mu_1)/c + m_1, \qquad c = \sqrt{\frac{m_2 - m_1^{\,2}}{\mu_2 - \mu_1^{\,2}}}$$

τότε έχουμε:

$$N^{-1}\sum_{i=1}^{N} y_i = m_1, \qquad N^{-1}\sum_{i=1}^{N} y_i^{\,2} = m_2$$

Αυτές οι μετασχηματισμένες ακολουθίες θα πρέπει να χρησιμοποιούνται με κάποια προσοχή. Αφού ο μετασχηματισμός εμπλέκει τις μ_1 και μ_2, τα νέα δείγματα-σημεία y_i δεν είναι πια ανεξάρτητα και οι εκτιμήσεις του σφάλματος της MC δεν είναι πια ξεκάθαρες. Για παράδειγμα το Κ. Ο. Θ μπορεί να μην εφαρμόζεται άμεσα. Στην πραγματικότητα, αυτή η μέθοδος είναι ένα παράδειγμα του δεύτερου τρόπου που χρησιμοποιούμε για να επιταχύνουμε την MC, με τροποποίηση των ιδιοτήτων της τυχαίας ακολουθίας, που θα δούμε στο επόμενο κεφάλαιο. Παρουσιάζεται όμως εδώ, γιατί η προκύπτουσα βελτίωση της ακρίβειας της MC είναι συγκρίσιμη με αυτήν που έχουμε χρησιμοποιώντας τις υπόλοιπες τεχνικές μείωσης της διασποράς.

5.4 Διαστρωματωμένη δειγματοληψία (Stratified Sampling)

Η διαστρωματωμένη δειγματοληψία συνδυάζει τα πλεονεκτήματα ενός πλέγματος με εκείνα των τυχαίων μεταβλητών. Ο κανόνας εδώ είναι να παίρνουμε δείγματα-σημεία όπου οι τιμές $f(x_i)$ είναι πιο πιθανό να δείξουν μεγάλη διασπορά. Στην πιο απλή περίπτωση, η διαστρωμάτωση βασίζεται σε ένα κανονικό πλέγμα με ομοιόμορφη πυκνότητα σε μια διάσταση. Διαμερίζουμε την περιοχή ολοκλήρωσης Ω=[0,1] σε *M* υποδιαστήματα Ω_k με κάθε ένα:

$$\Omega_k = \left[\frac{(k-1)}{M}, \frac{k}{M}\right]$$

Έτσι κάθε ένα τέτοιο υποσύνολο έχει το ίδιο μέτρο: $|\Omega_k| = \frac{1}{M}$. Ορίζουμε τις μέσες τιμές σε καθένα Ω_k με:

$$\bar{f}(x) = \bar{f}_k = |\Omega_k|^{-1} \int_{\Omega_\kappa} f(x)dx, \qquad \gamma\iota\alpha \ x \in \Omega_k$$

Για κάθε k, παίρνουμε $N_k = N/M$ σημεία $\{x_i^{(k)}\}$ ομοιόμορφα κατανεμημένα στο Ω_k.

Τοτε ο διαστρωματωμένος τύπος για την προσέγγιση με την MC είναι απλώς το άθροισμα των τύπων για κάθε υποδιάστημα, δηλαδή:

$$I_N = N^{-1} \sum_{k=1}^{M} \sum_{i=1}^{N/M} f\left(x_i^{(k)}\right)$$

Το σφάλμα της μεθόδου MC για αυτό το διαστρωματωμένο άθροισμα είναι:

$$\varepsilon \approx \mathrm{N}^{-1/2} \sigma_s$$

με $\sigma_s^{\ 2} = \int \left(f(x) - \bar{f}(x)\right)^2 dx = \sum_{k=1}^{M} \int_{\Omega_k} \left(f(x) - \bar{f}_k\right)^2 dx$

Για αυτόν τον διαστρωματωμένο κανόνα υπάρχει ένα πολύ απλό αποτέλεσμα. Η διαστρωματωμένη MC υπερνικάει πάντα την απλή MC μέθοδο, αφού:

$$\sigma_s \le \sigma$$

Η απόδειξη της ανισότητας αυτής είναι απλή (βλ. βιβλιογραφία Caflisch)

5.5 Σημαντική Δειγματοληψία (Importance Sampling)

Η σημαντική δειγματοληψία είναι η πλέον διαδεδομένη μέθοδος μείωσης της διασποράς. Στην απλή ολοκλήρωση Monte Carlo, χρησιμοποιήσαμε μια ομοιόμορφη κατανομή. Αυτή ήταν μόνο μια επιλογή. Εάν το ολοκλήρωμα που που υπολογίζουμε δεν είναι αρκετά ομοιόμορφο, ένας καλύτερος τρόπος για να μειώσουμε την διασπορά είναι να παίρνουμε περισσότερα δείγματα-αριθμούς

από κάποιες περιοχές παρά από κάποιες άλλες. Εάν δηλαδή η $f(x)$ διακυμαίνεται αρκετά στο πεδίο ολοκλήρωσης, υπάρχουν κάποια διαστήματα στα οποία η συνάρτηση είναι περισσότερο «σημαντική», έχει περισσότερο βάρος στο ολοκλήρωμα από ότι σε άλλα διαστήματα. Θα μπορούσαμε να χρησιμοποιήσουμε μια κατανομή με πεδίο ορισμού το πεδίο ολοκλήρωσης για να δώσουμε αξία σε αυτό το γεγονός, με την έννοια ότι θα παίρνουμε τυχαίους αριθμούς πιο συχνά από τα «σημαντικά» διαστήματα (με ένα τρόπο, έχουμε ήδη διερευνήσει αυτήν την μέθοδο στην διαστρωματωμένη δειγμαληψία).

Έστω ότι έχουμε μια τυχαία μεταβλητή ***X*** με κατανομή πουχαρακτηρίζεται από την συνάρτηση πυκνότητάς της $\phi(x)$ με πεδίο ορισμού το *D*. Τότε:

$$I = \int_{\mathbf{x}\in D} f(\mathbf{x})d\mathbf{x} = \int_{\mathbf{x}\in D} \frac{f(\mathbf{x})}{\phi(\mathbf{x})}\phi(\mathbf{x})d\mathbf{x} = E_X\left[\frac{f(\mathbf{x})}{\phi(\mathbf{x})}\right] \qquad (25)$$

Αυτή η μέση τιμή, και το ολοκλήρωμα, μπορεί εκτιμηθεί χρησιμοποιώντας την:

$$\hat{I}_n^{\phi} = \frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i)}{\phi(x_i)}$$

και η διασπορά και ο εκτιμητής της είναι:

$$V(\hat{I}) = \frac{1}{n}\left(\int_{x\in D}\frac{f(x)^2}{\phi(x)}dx - \left(\int_{x\in D} f(x)dx\right)^2\right)$$

$$\hat{V}(\hat{I}) = \frac{1}{n}\frac{1}{n-1}\left(\sum_{i=1}^{n}\frac{f(x_i)^2}{\phi(x_i)^2} - \hat{I}^2\right)$$

όπου τα x_i είναι δοκιμές από την κατανομή που χαρακτηρίζεται από την πυκνότητα $\phi(x)$. Είναι σαφές ότι η προς ολοκλήρωση $\frac{f(x)^2}{\phi(x)}$ είναι σημαντική, και εάν αυτή μπορούσε να ελαχιστοποιηθεί, θα παίρναμε μια μικρή διασπορά του εκτιμητή $\hat{I}_n^{\phi}$. Έτσι, καθώς μειώνουμε την διασπορά στην $\frac{f(x)}{\phi(x)}$, δηλαδή χρησιμοποιούμε μια $\phi(x)$ που είναι παρόμοια με την $f(x)$, μειώνουμε επίσης την διασπορά του εκτιμητή $\hat{I}$. Μια δυσκολία είναι ότι απαιτείται δειγματοληψία από την πυκνότητα p, αλλά αυτό μπορεί να γίνει με τη μέθοδο Acceptance-Rejection που παρουσιάσαμε σε προηγούμενο κεφάλαιο.

Παρατήρηση: *Σημειώνουμε ότι εάν η* $\frac{f(x)^2}{\phi(x)}$ *πηγαίνει στο άπειρο, όπου φ(x)→0 , η διασπορά γίνεται άπειρη. Είναι δηλαδή επικίνδυνο να διαλέγουμε συναρτήσεις πυκνότητας που προσεγγίζουν το μηδέν γρήγορα. Έτσι χρειάζεται να προσπαθήσουμε αρκετά για να το εμποδίσουμε αυτό με μια έξυπνη επιλογή της* $\phi(x)$*. Αυτό μπορεί εύκολα να γίνει εάν χρησιμοποιήσουμε την κανονική κατανομή σαν την κατανομή της σημαντικής δειγματοληψίας.*

Θα πρέπει τώρα να είναι σαφές ότι η έννοια “σημαντική δειγματοληψία» αναφέρεται στο γεγονός ότι δίνουμε έμφαση σε σπάνια αλλά σημαντικά γεγονότα, όπως είναι μικρές περιοχές στις οποίες η συνάρτηση $f(x)$ είναι μεγάλη.

Ενα Αριθμητικό Παράδειγμα

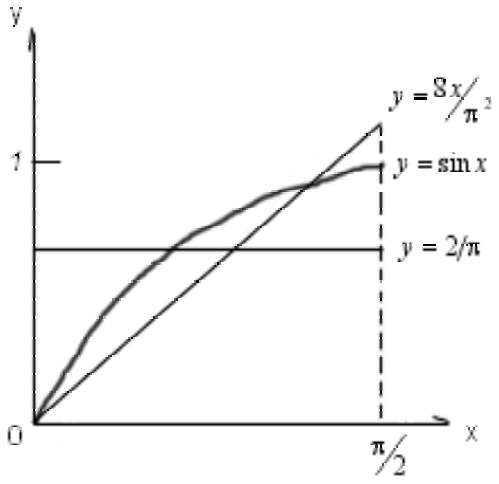
Θα προσπαθήσουμε να υπολογίσουμε το ολοκλήρωμα :

$$\mathbf{I} = \int_0^{\pi/2} \sin x \cdot dx$$

Η ακριβής τιμή του ολοκληρώματος αυτού είναι γνωστή και ίση με:

$$\int_0^{\pi/2} \sin x \cdot dx = \left[-\cos x\right]_0^{\pi/2} = 1$$

Θα χρησιμοποιήσουμε δύο διαφορετικές τυχαίες μεταβλητές X για τον υπολογισμό: μια με σταθερή πυκνότητα $p(x) = 2/\pi$, δηλαδή μια ομοιόμορφη κατανομή στο διάστημα *0 < x < π/2*, και μία τ.μ. με γραμμική πυκνότητα $p(x) = 8x/\pi^2$. Και οι δύο πυκνότητες, μαζί με την προς ολοκλήρωση *sinx* φαίνονται στο σχήμα.



Είναι προφανές ότι η γραμμική πυκνότητα συμφωνεί περισσότερο με την πρόταση της *δειγματοληπτικής σημασίας* , ότι είναι επιθυμητό η *p(x)* να είναι

ανάλογη με την απόλυτη τιμή της προς ολοκλήρωση συνάρτησης *sinx*. Έτσι μπορούμε να αναμένουμε ότι η δεύτερη προσέγγιση θα μας δώσει καλύτερα αποτελέσματα.

Πρώτη προσέγγιση

Έστω ότι $p(x) = 2/\pi$ για κάθε x στο (*0 , π/2).* Ο τύπος για την μοντελοποίηση της X προκύπτει από την σχέση x = *α +u (b-α)* για *α = 0* και *b = π/2,* όπου *u* είναι μια μεταβλητή ομοιόμορφη στο *(0,1)* οπότε θα έχουμε:

$$x = \frac{\pi}{2} u$$

Τώρα η εξίσωση της MC παίρνει τη μορφήQ

$$I \approx \frac{\pi}{2N} \sum_{j=1}^{N} \sin x_j$$

Θεωρούμε ότι *N* = 10. Για τις τιμές της *γ* χρησιμοποιούμε τριάδες ψηφίων από κάποιον πίνακα τυχαίων αριθμών πολλαπλασιασμένων με 0.001. Τα ενδιάμεσα αποτελέσματα φαίνονται στον παρακάτω πίνακα 1. Το τελικό αποτέλεσμα του υπολογισμού είναι $I \approx 0.952$

Δεύτερη προσέγγιση

Τώρα θεωρούμε ότι $p(x) = 8x/\pi^2$. Για τη μοντελοποίηση αυτής της τ.μ. αντιστρέφουμε την αθροιστική συνάρτηση κατανομής της

$$F(x) = P(X < x) = \int_0^x \frac{8t}{\pi^2} dt = \frac{4x^2}{\pi^2} = u$$

$$\Leftrightarrow$$

$$x = \frac{\pi}{2}\sqrt{u}$$

Τώρα η εκτίμηση MC παίρνει τη μορφή

$$I \approx \frac{\pi^2}{8N} \sum_{j=1}^{N} \frac{\sin x_j}{x_j}$$

Θεωρούμε πάλι *N* = 10. Χρησιμοποιούμε τους ίδιους τυχαίους αριθμούς *γ* όπως και στην πρώτη προσέγγιση. Τα ενδιάμεσα αποτελέσματα φαίνονται στον πίνακα 1.

Το τελικό αποτέλεσμα του υπολογισμού είναι:

$$I \approx 1.016$$

j	u_j	**Πρώτη προσέγγιση**		**Δεύτερη Προσέγγιση**	
		x_j	$\sin x_j$	x_j	$\frac{\sin x_j}{x_j}$
1	0.865	1.359	0.978	1.461	0.680
2	0.159	0.250	0.247	0.626	0.936
3	0.079	0.124	0.124	0.442	0.968
4	0.566	0.889	0.776	1.182	0.783
5	0.155	0.243	0.241	0.618	0.937
6	0.664	1.043	0.864	1.280	0.748
7	0.345	0.542	0.516	0.923	0.863
8	0.655	1.029	0.857	1.271	0.751
9	0.812	1.275	0.957	1.415	0.698
10	0.312	0.521	0.498	0.905	0.868

Πίνακας 1 : Ενδιάμεσα Αποτελέσματα των Δύο προσεγγίσεων

Όπως προβλέψαμε, η δεύτερη προσέγγιση δίδει περισσότερο ακριβές αποτέλεσμα.

Εκτιμήσεις Σφαλμάτων

Από τις τιμές που δίνονται στον πίνακα 1, μπορεί κανείς να προσεγγίσει τις διασπορές και για τις δύο περιπτώσεις. Για την πρώτη περίπτωση θα έχουμε:

$$\hat{V}ar\left(\frac{f(x)}{p(x)}\right) \approx \frac{1}{n-1}\left(\sum_{i=1}^{n}\frac{f(x_i)^2}{p(x_i)^2} - \hat{I}^2\right)$$

$$\approx \frac{\pi^2}{9\cdot 4}\left(\sum_{j=1}^{10}\left(\sin\xi_j\right)^2 - \frac{1}{10}\left(\sum_{j=1}^{10}\sin\xi_j\right)^2\right)$$

$$= \frac{\pi^2}{36}\left(4.604 - 3.670\right) = 0.256$$

Για τη δεύτερη περίπτωση θα έχουμε:

$$Var\left(\frac{f(x)}{p(x)}\right) \approx \frac{\pi^4}{9\cdot 64}\left(\sum_{j=1}^{10}\left(\frac{\sin\xi_j}{\xi_j}\right)^2 - \frac{1}{10}\left(\sum_{j=1}^{10}\frac{\sin\xi_j}{\xi_j}\right)^2\right) \Leftrightarrow$$

$$= \frac{\pi^4}{576}\left(6.875 - 6.777\right) = 0.016$$

Αν και το μέγεθος του δείγματος είναι μικρό (N=10) και δεν μπορεί να εγγυηθεί την εφαρμογή του Κεντρικού Οριακού Θεωρήματος, ας

προσπαθήσουμε να εκτιμήσουμε τα πιθανά σφάλματα $r_{\mathrm{N}} = 0.6745(Var()/N)^{1/2}$ και για τους δύο υπολογισμούς. Για την πρώτη περίπτωση, $r_N = 0.103$ ενώ για την δεύτερη $r_N = 0.027$. Φαίνεται καθαρά ότι τα πραγματικά απόλυτα σφάλματα - 0.048 και 0.016 αντίστοιχα - είναι της ίδιας τάξης μεγέθους με τα πιθανά σφάλματα.

Οι ακριβείς τιμές των διασπορών είναι 0.233 και 0.0166. Βλέπει κανείς οτι η δεύτερη προσέγγιση που περιλαμβάνει *σημαντική δειγματοληψία* είναι επίσης πιο ακριβής για την εκτίμηση της διασποράς.

5.6 Προσαρμόσιμες Monte Carlo – Μέθοδος Vegas

Οι τεχνικές μείωσης της διασποράς που περιγράφησαν παραπάνω απαιτούν κάποια a priori γνώση για την συμπεριφορά της συνάρτησης. Στην πράξη αυτή τη γνώση δεν την έχουμε συχνά, οπότε είναι καλύτερο να χρησιμοποιούμε προσαρμόσιμες τεχνικές, δηλαδή αλγορίθμους που μαθαίνουν για την συνάρτηση καθώς προχωρούν. Ένας τέτοιος αλγόριθμος είναι ο Vegas που χρησιμοποιείται πολύ στην Ενεργειακή Φυσική. Ο Vegas συγκεντρώνει τις βασικές ιδέες της σημαντικής δειγματοληψίας και της διαστρωμάτωσης σε έναν επαναληπτικό αλγόριθμο, που αυτόματα συγκεντρώνει τους συναρτησιακούς υπολογισμούς σε εκείνες τις περιοχές όπου η προς ολοκλήρωση συνάρτηση έχει μεγαλύτερο μέγεθος. Ο αλγόριθμος ξεκινά διαμερίζοντας το πεδίο ολοκλήρωσης σε ένα παραλληλόγραμμο πλέγμα και κάνει μια ολοκλήρωση σε κάθε υποπεδίο. Αυτά τα αποτελέσματα χρησιμοποιούνται κατόπιν για να προσαρμοστεί κατάλληλα το πλέγμα για την επόμενη επανάληψη, σύμφωνα με το από που το ολοκλήρωμα δέχεται κυρίαρχες επιδράσεις. Με αυτόν τον τρόπο, ο Vegas χρησιμοποιεί σημαντική δειγματοληψία και προσπαθεί να προσεγγίσει τη βέλτιστη συνάρτηση πυκνότητας πιθανότητας:

$$p_{OPTINAL}(x) = \frac{f(x)}{\int |f(x)| dx}$$

με μια συνάρτηση βήματος.

Τελικά μετά από κάποιες επαναλήψεις, βρίσκεται το καλύτερο πλέγμα. Μετά από αυτήν την αρχική εξερευνητική φάση, το πλέγμα παγώνει και σε μια δεύτερη φάση υπολογίζεται το ολοκλήρωμα με υψηλή ακρίβεια βάση του βέλτιστου αυτού πλέγματος. Ο διαχωρισμός αυτός σε δύο φάσεις μας επιτρέπει να χρησιμοποιούμε λιγότερους συναρτησιακούς υπολογισμούς στην εξερευνητική φάση και να αγνοούμε τις αριθμητικές εκτιμήσεις αυτής της φάσης που κατά κανόνα θα έχουν μεγαλύτερη διασπορά.

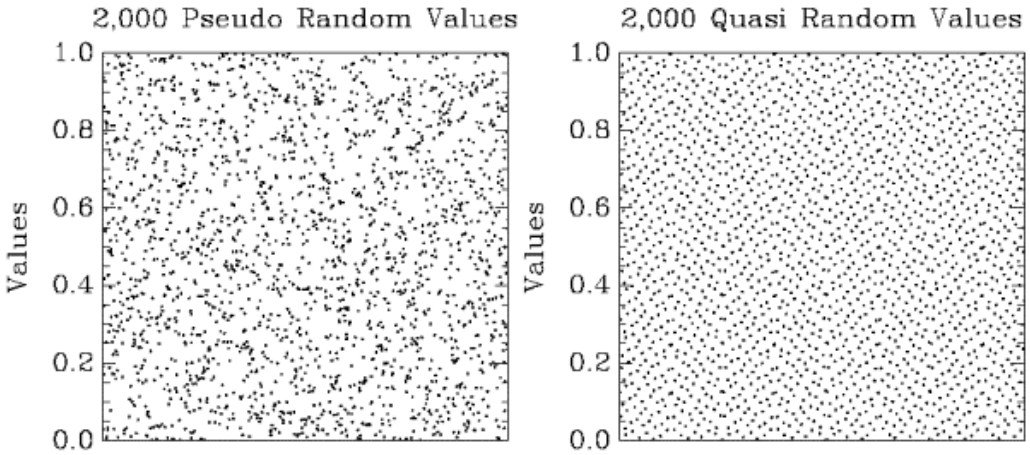
6. Quasi-Monte Carlo Μέθοδοι

Οι quasi-τυχαίες ακολουθίες είναι μια ντετερμινιστική εναλλακτική λύση για να επιταχύνουμε την MC. Σε αντίθεση με τις ψευδο-τυχαίες ακολουθίες αριθμών που είδαμε μέχρι τώρα, και οι οποίες προσπαθούν να μιμηθούν τις ιδιότητες των τυχαίων ακολουθιών, οι quasi-τυχαίες ακολουθίες είναι σχεδιασμένες έτσι ώστε να μας παρέχουν μεγαλύτερη ομοιομορφία από μια τυχαία ακολουθία και έτσι γρηγορότερη σύγκλιση στους αριθμητικούς τύπους προσέγγισης. Η ομοιομορφία μιας ακολουθίας μετράται σε σχέση με την ασυμφωνία (discrepancy) της, και γι΄αυτό πολλές φορές οι quasi-τυχαίες ακολουθίες καλούνται και χαμηλής-ασυμφωνίας (low discrepancy). Οι μέθοδοι που προκύπτουν με την χρήση των ακολουθιών αυτών λέγονται quasi-Monte Carlo.

6.1 Monte Carlo εναντίον quasi-Monte Carlo

Οι τυπικές μέθοδοι Monte Carlo χρησιμοποιούν (ψευδο-)τυχαίες ακολουθίες αριθμών για συγκλίνουν όπως έχουμε δει με ρυθμό $O\left(N^{-1/2}\right)$, όπου N ο αριθμός των δειγμάτων-σημείων.Επιπλέον οι κλασσικές MC μέθοδοι χρησιμοποιούνται εκτός από την ολοκλήρωση, και στην βελτιστοποίηση αλλά και στα προβλήματα προσομείωσης.

Ο περιοριστικός παράγοντας στην ακρίβεια της MC είναι ο συστοιχισμός (clumbing) που εμφανίζεται στα σημεία μιας (ψευδο-)τυχαίας ακολουθίας. Ο συστοιχισμός αυτός όπως επίσης και τα κενά χωρίς καθόλου σημεία, που φαίνονται καθαρά στο παρακάτω σχήμα, όπου συγκρίνουμε οπτικά μια (ψευδο-) τυχαία ακολουθία με μια quasi-τυχαία ακολουθία (Sobol') στις δύο διαστάσεις.



Ο λόγος που η (ψευδο)τυχαία ακολουθία (αριστερό σχήμα) εμφανίζει αυτούς τους συστοιχισμούς αλλά και τα κενά είναι ότι τα σημεία είναι ανεξάρτητα μεταξύ τους. Αφού διαφορετικά σημεία δεν γνωρίζουν τίποτε το ένα για το άλλο, υπάρχει μια μικρή πιθανότητα να πέσουν πολύ κοντά μεταξύ τους. Στην πράξη, περίπου

$\sqrt{N}$ σημεία από τα *N* συνολικά πέφτουν σε τέτοιες συστοιχίες.

Οι quasi-Monte Carlo μέθοδοι χρησιμοποιούν quasi-τυχαίους αριθμούς, που είναι ντετερμιστικοί, με συσχετίσεις μεταξύ των σημείων για να αποφευχθούν οι συστοιχίες αυτές. Έτσι οι quasi Monte Carlo μέθοδοι έχουν ρυθμό σύγκλισης $O\left((\log N)^k N^{-1}\right)$. Λόγω των συσχετίσεων αυτών δεν είναι τόσο ευπροσάρμοστες, πολλαπλών χρήσεων όπως η κλασσική Monte Carlo με τους ψευδο-τυχαίους αριθμούς. Είναι σχεδιασμένοι για ολοκλήρωση παρά για προσομείωση ή βελτιστοποίηση. Βέβαια, το ζητούμενο αποτέλεσμα μιας προσομείωσης μπορεί συχνά να γραφεί σαν μια εκτίμηση, που είναι ένα ολοκλήρωμα, έτσι ώστε να μπορεί να εφαρμοστεί μια μέθοδος qMC. Όπως θα δούμε όμως παρακάτω η χρήση quasi-τυχαίων ακολουθιών εισάγει υψηλή διαστατικότητα, που μπορεί να περιορίσει την αποτελεσματικότητα των quasi-Monte Carlo ακολουθιών.

6.2 Ασυμφωνία (discrepancy)

Οι quasi-τυχαίες ακολουθίες εφευρέθηκαν από τους θεωρητικούς των αριθμών που ενδιαφέρονταν για τις ιδιότητες ομοιομορφίας των αριθμητικών ακολουθιών (Kuipers και Niederreiter 1974, Hua και Wang 1981). Ένα σημαντικό πρώτο βήμα είναι η τυποποίηση ενός ποσοτικού μέτρου ομοιομορφίας. Η ομοιομορφία μιας ακολουθίας μετράται με την *ασυμφωνία* (discrepancy) της.
Ας το δούμε λιγάκι διαισθητικά πρώτα: Φανταζόμαστε ότι έχουμε 1000 σημεία σε ένα κουτί που είναι ένας κύβος με όγκο 1. Πως θα κατανείμουμε τα 1000 αυτά σημεία στο χώρο του κύβου έτσι ώστε οποιονδήποτε υπο-χώρο πάρουμε στο κουτί να έχει έναν αριθμό σημείων μέσα του που να είναι ανάλογος με το αντίστοιχο όγκο του; Αυτό είναι ένα κλασσικό πρόβλημα θεωρίας μέτρου που λύθηκε από τον μαθηματικό Roth. Αυτός τυποποίησε το πρόβλημα της χαμηλής ασυμφωνίας (low discrepancy). Η ασυμφωνία συσχετίζεται με την χειρότερη απόκλιση του ποσοστού των σημείων σε κάθε υποόγκο (ή υπο – τετράγωνο αν φανταστομε μια λεπτή φέτα στο χώρο) σε σχέση με το ακριβές ποσοστό των σημείων για τέλεια ομοιόμορφα κατανεμημένα σημεία. Δουλεύοντας με τον μοναδιαίο υπερκύβο, εργαζόμαστε απευθείας με το ποσοστό του όγκου του για τους υποόγκους.
Η ασυμφωνία μετράει τον βαθμό στον οποίον τα σημεία είναι εξίσου διασκορπισμένα στον μοναδιαίο υπερκύβο. Έστω ένας d-διάστατος υπερκύβος $I^d = [0,1)^d$ με όγκο 1 και μια ακολουθία P, $\{\mathbf{x}_n\}$ N σημείων διασκορπισμένων σε όλον όγκο του υπερκύβου. Έστω J να είναι ένα υποτετράγωνο, δηλαδή μια λεπτή φέτα του υπερκύβου I^d , με όγκο v(J) και E να είναι το σύνολο όλων των τέτοιων συνόλων-υποτετραγώνων του υπερκύβου. Τότε ορίζουμε :

$$R_N(J) = \frac{\textit{αριθμός σημείων στο } J}{N} - v(J)$$

Το $R_N(J)$ είναι απλώς το σφάλμα της Monte Carlo στην μέτρηση του όγκου του J. Η ασυμφωνία ορίζεται τότε εφαρμόζοντας κάποια νόρμα στο $R_N(J)$.

Εάν εφαρμόσουμε την L-άπειρο νόρμα έχουμε την εξής ασυμφωνία D_N της P:

$$D_N = \sup_{J \in E} \left| R_N(J) \right|$$

Αυτή η έννοια συνδέεται με την χειρότερη περίπτωση ανάμεσα σε όλες για τα κενά και τις συστοιχίες σημείων.
Επιπλέον, επειδή κάθε τετραγωνικό σύνολο σαν το *J* μπορεί να οριστεί από δύο κορυφές-αντίποδες, ένα τέτοιο σύνολο *J* μπορεί να οριστεί σαν ***J=J(x,y)***, στο οποίο τα σημεία **x** και **y** είναι αντίποδες κορυφές. Έτσι, ορίζουμε και μια ακόμη ασυμφωνία, T_N, αν εφαρμόσουμε την L-τετράγωνο νόρμα στο $R_N(J)$:

$$T_N = \left[\int_{(x,y) \in I^{2d}, x_i < y_i} R_N\left(J(\mathbf{x},\mathbf{y})\right)^2 d\mathbf{x} d\mathbf{y} \right]^{1/2}$$

Επίσης ορίζουμε την ασυμφωνία για τετραγωνικά σύνολα με μια κορυφή 0, *J(0,**x***) δηλαδή σύνολα της μορφής: $\prod_{i=1}^{d} [0, x_i]$, και συμβολίζουμε E^* το σύνολο όλων των τέτοιων υποσυνόλων. Τότε η προκύπτουσα ασυμφωνία, ασυμφωνία άστρο, είναι για τις δύο παραπάνω περιπτώσεις:

$$D_N^* = \sup_{J \in E^*} \left[R_N(J) \right]$$

και

$$T_N^* = \left[\int_{I^d{}_i} R_N\left(J(0,\mathbf{x})\right)^2 d\mathbf{x} \right]^{1/2}$$

6.3 Koksma-Hlawka ανισότητα

Οι quasi-τυχαίες ακολουθίες είναι χρήσιμες στην ολοκλήρωση γιατί οδηγούν σε πολύ μικρότερο σφάλμα από ότι η κλασσική Monte Carlo με τους (ψευδο-)τυχαίους αριθμούς. Η βάση για την ανάλυση του σφάλματος σύγκλισης στις quasi-Monte Carlo μεθόδους είναι η ανισότητα Koksma-Hlawka.
Θεωρούμε το ολοκλήρωμα

$$I[f] = \int_{I^d} f(x) dx$$

και την προσέγγιση Monte Carlo:

$$I_N[f] = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

και έχουμε ορίσει το σφάλμα σύγκλισης:

$$\varepsilon[f] = I\,[f] - I_N[f]$$

Ορίζουμε την μεταβλητότητα (variation) (με την Hardy Krause έννοια) της συνάρτησης f, μιας μεταβλητής:

$$V[f] = \int_0^1 \left| \frac{df}{dt} \right| dt$$

Σε d διαστάσεις η μεταβλητότητα ορίζεται ως:

$$V[f] = \int_{I^d} \left| \frac{\partial^d f}{\partial t_1 \ldots \partial t_d} \right| dt_1 \ldots dt_d + \sum_{i=1}^{d} V\left| f_1^{(i)} \right|$$

όπου $f_1^{(i)}$ είναι ο περιορισμός της συνάρτησης f στο σύνορο $x_i = 1$. Αφού αυτοί οι περιορισμοί είναι συναρτήσεις d-1 μεταβλητών , ο ορισμός αυτός είναι επαναληπτικός.

Θεώρημα Koksma-Hlawka

Για κάθε ακολουθία $\{\mathbf{x}_n\}$ και για κάθε συνάρτηση f με φραγμένη μεταβλητότητα, το σφάλμα ολοκλήρωσης ε φράσσεται ως εξής:

$$\varepsilon[f] \le V[f] \cdot D_N^*$$

Παρατηρούμε ότι η ακρίβεια ελέγχεται από τη μεταβλητότητα V και από την ασυμφωνία της ακολουθίας. Επειδή συνήθως δεν διαλέγουμε εμείς την συνάρτηση προς ολοκλήρωση και άρα δεν μπορούμε να ελέγξουμε την μεταβλητότητα της, προσπαθούμε να κατασκευάσουμε ακολουθίες χαμηλής ασυμφωνίας σε μεγάλες διαστάσεις για αποτελεσματική ολοκλήρωση.

Είναι χρήσιμο να συγκρίνουμε την Koksma – Hlawka ανισότητα για τις quasi-MC με το μέσο τετραγωνικό σφάλμα (Root Mean Square Error (RMSE)) της κλασσικής MC ολοκλήρωσης που χρησιμοποιεί (ψευδο-)τυχαίες ακολουθίες:

$$E\left[\varepsilon[f]^2\right]^{1/2} = \sigma[f] \cdot N^{-1/2}$$

όπου :

$$\sigma[f] = \left(\int_{\mathrm{I}^d} \left(f(x) - I[f] \right)^2 dx \right)^{1/2}$$

και

$$\varepsilon[f] = \int_{I^d} f(x)dx - \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

Παρατηρούμε τα εξής :

- Και τα δύο φράγματα του σφάλματος είναι ένα γινόμενο ενός παράγοντα που εξαρτάται από την ακολουθία (το $N^{-1/2}$ για το RMSE και η D_N για την Koksma-Hlawka) και ενός που εξαρτάται από την συνάρτηση (η μεταβλητότητα V για την Koksma-Hlawka και η διασπορά σ για το RMSE)
- Η ανισότητα Koksma-Hlawka είναι ένα φράγμα χειρότερης περίπτωσης, ενώ το μέσο τετραγωνικό σφάλμα της κλασσικής MC είναι πιθανοθεωρητικό.
- Το RMSE είναι ισότητα και για αυτό το λόγο στιβαρό, ενώ η ανισότητα είναι απλώς ένα άνω φράγμα.
- Η εμπειρία μας δείχνει ότι η μεταβλητότητα V στην Koksma-Hlawka είναι συνήθως μια μεγάλη υπερεκτίμηση, ενώ η ασυμφωνία είναι ενδεικτική της πραγματικής απόδοσης.

Ένα επιπλέον, πιο περίπλοκο σημείο είναι ότι στην κλασσική Monte Carlo, κάθε σημείο είναι μια εκτίμηση του όλου ολοκληρώματος. Σε μια τυπική quasi-τυχαία ακολουθία, από την άλλη, έχουμε μια ιεραρχική δομή, έτσι ώστε τα αρχικά σημεία να δειγματίζουν το ολοκλήρωμα σε μια αρχική κλίμακα και τα κατοπινά σημεία να το προσεγγίζουν όλο και καλύτερα. Αυτή είναι η πηγή των logN όρων στα φράγματα της ασυμφωνίας που θα δούμε παρακάτω.

6.4 Μέσο Σφάλμα Ολοκλήρωσης

Η περισσότερη υπολογιστική εμπειρία μας δείχνει ότι η ασυμφωνία είναι ένας καλός δείκτης για την απόδοση της quasi-Monte Carlo, ενώ η μεταβλητότητα δεν είναι.
Αυτή ακριβώς η εμπειρία επιβεβαιώνεται και από το εξής αποτέλεσμα που απέδειξε το 1990 ο Wozniakowski

Θεώρημα Wozniakowski

Έχουμε ότι: $E\left[\varepsilon_N[f]^2\right]^{1/2} = T_N^*$,

Οπου η μέση τιμή *E* λαμβάνεται σε σχέση με τις συναρτήσεις *f που είναι κατανεμημένες σύμφωνα με Brownian μέτρο,* το οποίο είναι ένα μέτρο του χώρου των συναρτήσεων: με πιθανότητα 1, μια τέτοια συνάρτηση $f(\mathbf{x})$ είναι «ημι-διαφορίσιμη». Ειδικά, η μεταβλητότητά της είναι άπειρη. Σε αυτό το πλαίσιο μπορούμε να πούμε με βεβαιότητα ότι η μεταβλητότητα δεν είναι καλός δείκτης του σφάλματος της ολοκλήρωσης. Αφού η ταυτότητα στο θεώρημα του

Wozniakowski είναι μια ισότητα, έπεται ότι η L-τετράγωνο ασυμφωνία T_N^* συμφωνεί όντως με το τυπικό μέγεθος του σφάλματος στην quasi-Monte Carlo ολοκλήρωση. Επιπλέον, η υπολογιστική εμπειρία μας δείχνει ότι συνήθως οι ασυμφωνίες T_N^* και D_N είναι του ίδιου μεγέθους.
Γι' αυτούς τους λόγους, προσπαθούμε να κατασκευάσουμε ακολουθίες τυχαίων αριθμών με χαμηλή ασυμφωνία.

6.5 Βασικές Γεννήτριες quasi-τυχαίων αριθμών

Μια άπειρη ακολουθία $\{\mathbf{x}_n\}$ είναι *ομοιόμορφα κατανεμημένη* αν:

$$\lim_{N\to\infty} D_N = 0$$

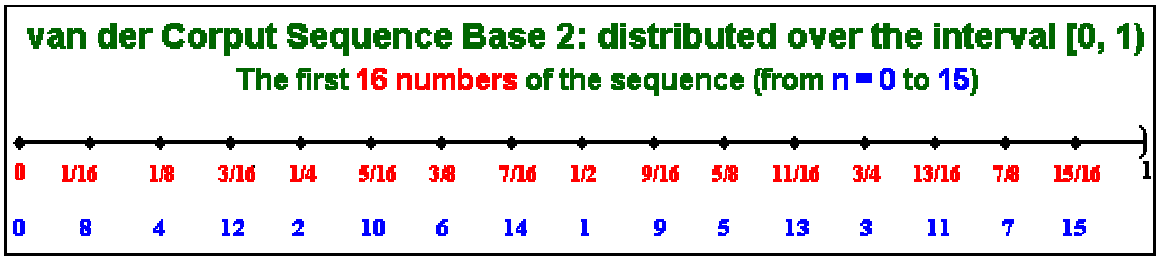
Μια ακολουθία λέγεται quasi-τυχαία ακολουθία αν:

$$D_N \le c(\log N)^k N^{-1}$$

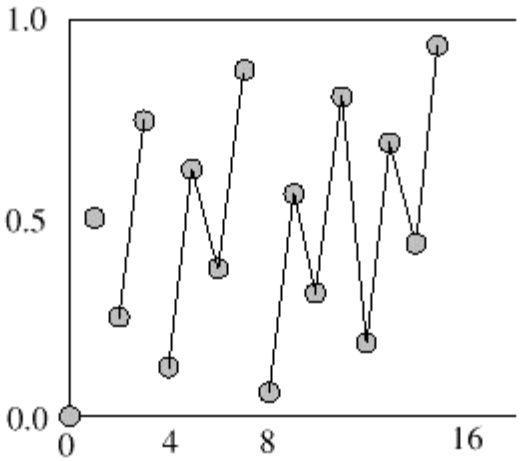
όπου c και k είναι σταθερές ανεξάρτητες του αριθμού των δειγμάτων *N*, αλλά μπορεί να εξαρτώνται από τη διαστάση d. Ιδιαίτερα, είναι δυνατό να κατασκευάσουμε ακολουθίες για τις οποίες ισχύει : *k = d.* Είναι σύνηθες επίσης να λέμε ότι μια ακολουθία είναι quasi-τυχαία μόνο αν ο εκθέτης k είναι ίσος με τη διάσταση d της ακολουθίας.
Σε πρακτικούς υπολογισμούς, δεν ενδιαφερόμαστε τόσο να βρούμε μια πεπερασμένη ακολουθία $x_1, x_2, ..., x_N$ που πετυχαίνει την μικρότερη ασυμφωνία για κάποιο δεδομένο N, αλλά μια άπειρη ακολουθία που τα πρώτα *N* στοιχεία της έχουν μικρή ασυμφωνία. Τα πλεονεκτήματα είναι ότι μπορούμε να αλλάξουμε την τιμή του *N* σε μια quasi-Monte Carlo ολοκλήρωση χωρίς να χάσουμε τις προηγούμενα υπολογισμένες συναρτησιακές τιμές. Στην πράξη, χρησιμοποιούμε συγκεκριμένες ακολουθίες για τις οποίες γνωρίζουμε ότι η ασυμφωνία είναι μικρότερη από $D_N \le c(\log N)^k N^{-1}$ για κάποιο συγκεκριμένο k. Δηλαδή, πρέπει να τονιστεί ότι οι καλά-κατανεμημένες ακολουθίες είναι σχεδιασμένες για να είναι καλά κατανεμημένες για κάποιο συγκεκριμένο αριθμό διαστάσεων και όχι για κάθε αριθμό διαστάσεων.

Η βασική ιδέα της χαμηλής ασυμφωνίας είναι συνδεδεμένη με την ιδιότητα ότι οι διαδοχικοί αριθμοί προστίθενται σε μια θέση όσο πιο μακρυά γίνεται από τους προηγούμενους αριθμούς της ακολουθίας, δηλαδή αποφεύγοντας τον συστοιχισμό (ομάδες αριθμών κοντά μεταξύ τους - clustering). Η ακολουθία έτσι κατασκευάζεται βάση ενός σχήματος στο οποίο κάθε σημείο απωθείται από τα άλλα. Έτσι αν η ιδέα για τα σημεία είναι να αποφεύγουν όσο το δυνατόν τα υπόλοιπα, αυτό που πρέπει να κάνουν οι αριθμοί που παράγονται διαδοχικά είναι να καλύπτουν τα κενά που άφησαν οι προηγούμενοι αριθμοί μεταξύ τους.

Η επόμενη εικόνα δείχνει ακριβώς αυτή τη δυναμική γεμίσματος των κενών από την quasi-ακολουθία Van der Corput (1935) σε βάση 2, η οποία είναι ο θεμέλιος λίθος πολλών σημαντικών ακολουθιών χαμηλής ασυμφωνίας. Η ακολουθία ξεκινά από το 0 και περιορίζεται στο διάστημα [0,1) με τους πρώτους 16 (n=0 …15) αριθμούς να δίνονται από :



Χρησιμοποιούμε το διάστημα [0,1) για τον ορισμό της ακολουθίας, κλειστό στο 0, γιατί από εκεί ξεκινά η ακολουθία, αλλά ανοικτό στο 1, γιατί η ακολουθία ποτέ δεν φτάνει τον αριθμό 1, αν και για πολύ μεγάλο n μπορούμε να πάρουμε αριθμούς οσοδήποτε κοντά στο 1 θέλουμε.
Όπως φαίνεται στην παραπάνω φωτογραφία δεν υπάρχουν συστοιχισμοί σημείων και αυτό είναι ακριβώς ότι θέλαμε να πετύχουμε.
Στο παρακάτω σχήμα φαίνεται καθαρά ότι η ακολουθία Van der Corput επιδεικνύει μια κυκλικότητα που μπορεί να είναι πρόβλημα για πολλές εφαρμογές: πηγαίνει προς το 1 και γυρνάει πίσω σε κύκλους δύο σημείων. Πρακτικά μπορούμε με μια τυχαία μετάθεση της ακολουθίας να εξαλείψουμε αυτήν την εξάρτηση.



Αυτή δεν είναι η μόνη δυνατή ακολουθία Van der Corput. Μπορούμε να κατασκευάσουμε άλλες ακολουθίες αλλάζοντας βάση σε κάποιον άλλο πρώτο αριθμό. Διαφορετικές βάσεις μας δίνουν διαφορετικό μήκος κύκλων, πλήθος αριθμών που καλύπτουν το διάστημα [0,1) σε κάθε κύκλο. Στην βάση 2, τα ζεύγη (0, 1/2) και (1/4, 3/4) είναι τα πρώτα δύο ζεύγη. Στην βάση 3, για παράδειγμα, η ακολουθία έχει μήκος κύκλων 3, και είναι πάει ως εξής:

0, 1/3 , 2/3 , 1/9 , 4/9 , 7/9 , 2/9 , 5/9 , 8/9 , 1/27 , 10/27, 19/27, …

όπου ο πρώτος κύκλος είναι (0, 1/3, 2/3).

Η παραγωγή της ακολουθίας Van der Corput, για το n-ιοστό στοιχείο και σε βάση b, γίνεται με την εξής διαδικασία 3 βημάτων:

I. Ο αριθμός n με βάση το 10, αναπτύσσεται σε βάση b. Για παράδειγμα το τέταρτο στοιχείο (n=4) σε βάση 2 θα γίνει: $4_{10} = 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 = 100_2$

II. Ο αριθμός που πήραμε για τη βάση b, κατοπτρίζεται γύρω από τη δεκαδική τελεία. Στο παραπάνω παράδειγμά μας παίρνουμε 001 → 0.001

III. Παίρνουμε τον αριθμό σε βάση 10 που αντιστοιχεί στον αριθμό σε βάση 2: Στο παράδειγμά μας ο αριθμός 0.001 βάσης 2, αντιστοιχεί στον δεκαδικό 1/8 , αφού $0.001_2 = 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 1/8_{10} = 0.125_{10}$. Άρα ο τέταρτος αριθμός της ακολουθίας μας είναι ο 0.125.

Τα παραπάνω βήματα μπορούν να συνοψιστούν σε δύο εξισώσεις. Στην πρώτη παίρνουμε τα αντιστραμμένα ψηφία a_j του n στη βάση b, ενώ στην δεύτερη εξίσωση παίρνουμε τον αντίστοιχο του δεκαδικού n αριθμό της ακολουθίας Van der Corput:

$$n = \sum_{j=0}^{m} a_j(n)\, b^j$$

$$b(n) = \Phi_b(n) = \sum_{j=9}^{m} a_j(n) b^{-j-1}$$

όπου τα $a_j(n) \in \{0,1,...,b-1\}$ και πεπερασμένος αριθμός από αυτά είναι μηδεν.
Έτσι παίρνουμε ότι ο n-ιοστός αριθμός της ακολουθίας θα δίνεται από :

$$x_n = \Phi_b(n) \quad \text{n=0,1….}$$

Προφανώς κάθε φορά που το N είναι δύναμη του 2, τα σημεία $x_0, x_1, ..., x_{N-1}$ απέχουν εξίσου στο [0,1) και έτσι έχουν την μικρότερη ασυμφωνία. Εν γένει, η ασυμφωνία της ακολουθίας Van der Corput ικανοποιεί την :

$$D_N^*(S) \sim O\left(N^{-1} \log N\right)$$

Μια γενίκευση είναι οι ακολουθίες Halton. Στις k διαστάσεις, παίρνουμε μια ακολουθία βάσεων $p_1, p_2, ..., p_k$ πρώτων αριθμών, συνήθως οι πρώτοι k πρώτοι. Χρησιμοποιούμε μετά διαφορετική βάση για κάθε στοιχείο του αριθμού Halton x_n. Έχουμε δηλαδή:

$$x_n = \left(\Phi_{p_1}(n), \Phi_{p_2}(n), ..., \Phi_{p_k}(n)\right)$$

Οπότε για να πάρουμε κάθε ένα στοιχείο του x_n, αναπτύσσουμε κάθε φορά στην βάση p_i, στον ι-στο πρώτο, και αντιστρέφουμε γύρω από τη δεκαδική τελεία.
Η ασυμφωνία των ακολουθιών Halton είναι φραγμένη:

$$D_N(Halton) \le c_d (\log N)^d N^{-1}$$

όπου η σταθερά c εξαρτάται από την διάσταση d.

Επιπλέον quasi-τυχαίες ακολουθίες υπάρχουν:
Haselgrove (1961)
Faure (1982)
Sobol' (1967, 1976)
Niederreiter (1992)
Owen (1995)

Για όλες αυτές τις ακολουθίες με ασυμφωνία της τάξης του $O\left((\log N)^d N^{-1}\right)$, η ανισότητα Koksma-Hlawka συνεπάγεται ότι το σφάλμα ολοκλήρωσης είναι του ίδιου μεγέθους, δηλαδή:

$$\varepsilon[f] \le c V[f] (\log N)^d N^{-1}$$

6.6 Περιορισμοί

Η ανισότητα Koksma-Hlawka και τα φράγματα της ασυμφωνίας για μια quasi-τυχαία ακολουθία συνεπάγονται ότι η σύγκλιση της quasi-Monte Carlo είναι πιο γρήγορη από την κλασσική MC. Από την άλλη πλευρά υπάρχουν αρκετοί περιορισμοί στην αποτελεσματικότητα των qMC μεθόδων.
Πρωτον, δεν υπάρχει θεωρητική βάση για εμπειρικές εκτιμήσεις της ακρίβειας των quasiMC μεθόδων. Ενώ όπως είδαμε προηγούμενα το Κεντρικό Οριακό Θεώρημα μπορεί να χρησιμοποιηθεί για να ελέγξουμε την ακρίβεια της κλασσικής Monte Carlo και για να προβλέψουμε τον απαιτούμενο αριθμό N δειγμάτων για μια συγκεκριμένη ακρίβεια. Αφού δεν υπάρχει ΚΟΘ για τις quasi-Monte Carlo ακολουθίες δεν υπάρχει αντίστοιχα εμπειρική εκτίμηση σφάλματος για την quasi-Monte Carlo. Από την άλλη πλευρά, η εμπιστοσύνη στην ακρίβεια της quasi-MC ολοκλήρωσης προκύπτει από την εκλέπτυνση ενός συνόλου υπολογισμών.
Δεύτερον, οι quasi-Monte Carlo μέθοδοι έχουν σχεδιαστεί για την ολοκλήρωση και δεν είναι απευθείας εφαρμόσιμες στην προσομείωση γενικά. Αυτό οφείλεται στις συσχετίσεις μεταξύ των σημείων μιας quasi-τυχαίας ακολουθίας. Όμως, σε πολλές προσομειώσεις το επιθυμητό αποτέλεσμα είναι μια μέση τιμή μιας ποσότητας, που μπορεί να γραφτεί σαν ένα πολυδιάστατο ολοκλήρωμα στο οποίο η quasi-Monte Carlo μπορεί να χρησιμοποιηθεί. Πράγματι, μπορούμε να θεωρούμε τις διαφορετικές διαστάσεις ενός quasi-τυχαίου σημείου σαν ανεξάρτητες μεταξύ τους, έτσι ώστε οι quasi-τυχαίες ακολουθίες να μπορούν να αναπαριστούν μια προσομείωση διανέμοντας μια διάσταση σε κάθε βήμα χρόνου

ή απόφαση. Αυτή η προσέγγιση συχνά απαιτεί μια υψηλής διάσταση quasi-τυχαία ακολουθία.
Τρίτον, η quasi-Monte Carlo ολοκλήρωση φαίνεται να χάνει την αποτελεσματικότητά της όταν η διάσταση του ολοκληρώματος γίνεται μεγάλη. Αυτό μπορεί να προεξοφληθεί από το φράγμα $O\left((\log N)^d N^{-1}\right)$ της ασυμφωνίας. Για μεγάλο d, πολλές διαστάσεις, αυτό το φράγμα κυριαρχείται από τον λογαριθμικό όρο εκτός αν:

$$N > 2^d$$

Τέταρτον, η quasi-Monte Carlo ολοκλήρωση φαίνεται να χάνει επίσης την αποτελεσματικότητά της αν η προς ολοκλήρωση συνάρτηση δεν είναι λεία. Ο παράγοντας της μεταβλητότητας *V[f]* στην ανισότητα Koksma-Hlawka είναι ένας δείκτης αυτής της εξάρτησης, αν και στην πραγματικότητα βρίσκουμε ότι αν η συνάρτηση είναι έστω και λίγο λεια, κάπου μεταξύ της συνέχειας και διαφορισιμότητας, είναι αρκετό. Αυτός ο περιορισμός είναι σημαντικός, αφού πολλές μέθοδοι Monte Carlo εμπλέκουν αποφάσεις κάποιου είδους, που συνήθως αντιστοιχούν σε πολλαπλασιασμό με 0 ή 1.
Τελικά, ένα σημαντικό εμπειρικό συμπέρασμα είναι ότι οι quasi-Monte Carlo μέθοδοι ολοκλήρωσης είναι σχεδόν τόσο ακριβείς όσο η κλασσική Monte Carlo ολοκλήρωση. Έτσι η «απώλεια της αποτελεσματικότητας» παραπάνω σημαίνει πρακτικά ότι η quasi-Monte Carlo λειτουργεί όχι καλύτερα από την κλασσική Monte Carlo.
Στον παρακάτω πίνακα, έχουμε μια σύγκριση των quasi-MC και της κλασσικής Monte Carlo, για τον αριθμό των διαστάσεων d και τον αριθμό των δειγμάτων N, όπου στις quasi-Monte Carlo δίνεται τόσο η βέλτιστη περίπτωση όπου το σφάλμα ολοκλήρωσης είναι της τάξης του 1/N όσο και η χειρότερη περίπτωση όπου το σφάλμα έχει την τιμή του άνω φράγματος:

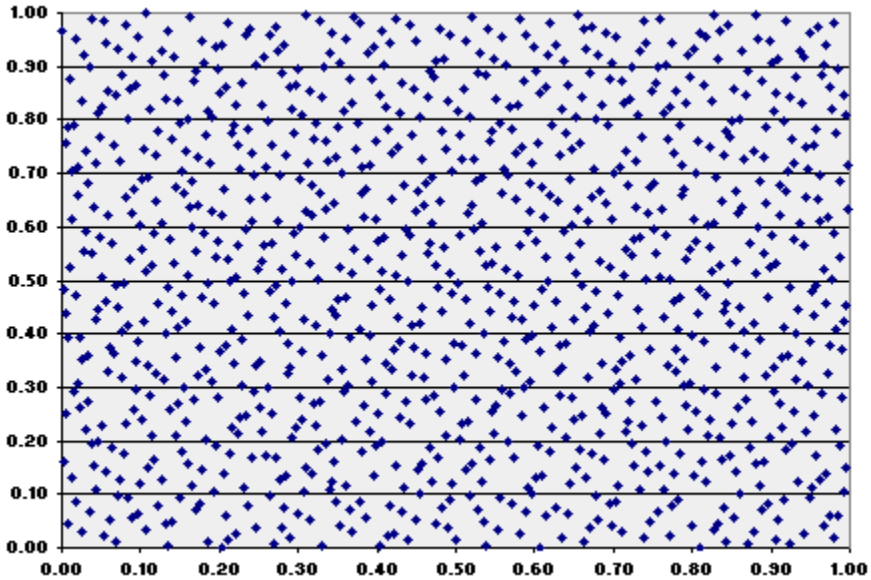
		Monte Carlo	Quasi-Monte Carlo	
			Best	Worse
d	N	$1/\sqrt{N}$	1/N	$\ln(N)^d/N$
1	1,000	0.03162	0.00100	0.00691
1	100,000	0.00316	0.00001	0.00012
2	10,000	0.01000	0.00010	0.00848
5	10,000	0.01000	0.00010	6.628
10	10,000	0.01000	0.00010	439295.5
50	100,000	0.00316	0.00001	1.14626E+48

OBS: without reduction of variance techniques for both MC and QMC

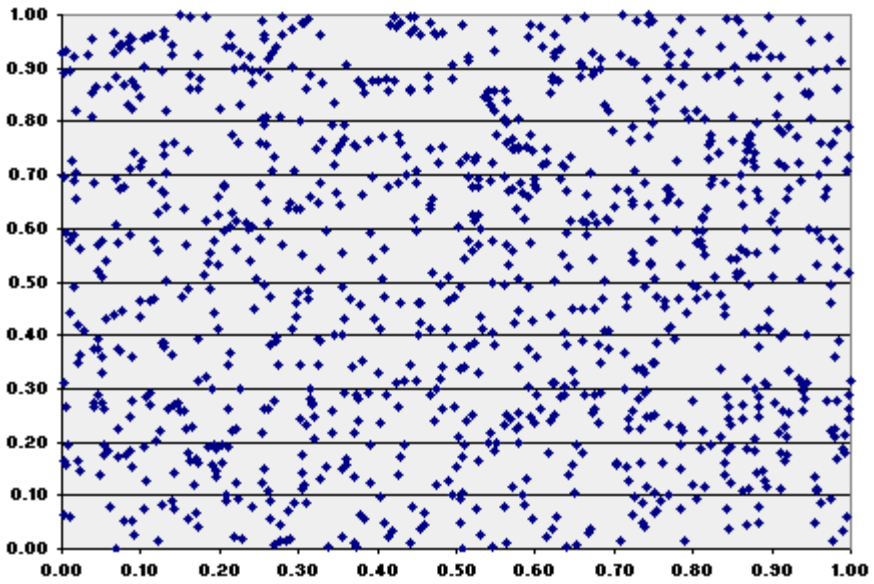
Επιπλέον, πολλές αλλά όχι όλες, οι quasi ακολουθίες έχουν την εξής ιδιότητα. Όλες οι μονοδιάστατες προβολές της ακολουθίας, δηλαδή όλα τα μονά στοιχεία, είναι πολύ καλά εξίσου κατανεμημένα. Αυτό συνεπάγεται ότι κάθε συνάρτηση που αποτελείται από άθροισμα μονοδιάστατων συναρτήσεων θα ολοκληρωθεί

πολύ καλά από μια τέτοια quasi-Monte Carlo μέθοδο. Ιδιαίτερα, οι quasi-Monte Carlo μέθοδοι δουλέυουν πολύ καλά για γραμμικές συναρτήσεις: $f(\mathbf{x}) \approx \sum_{i}^{d} a_i x_i$.

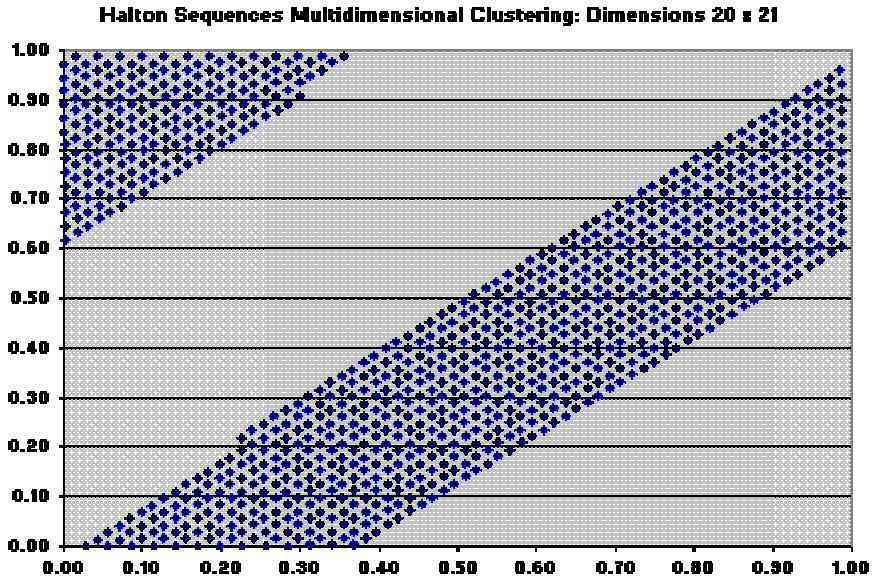
Τώρα ας θεωρήσουμε τις δυσδιάστατες αναπαραστάσεις. Για μια οπτική αναπαράσταση της αναποτελεσματικότητας των quasi-τυχαίων ακολουθιών όταν οι διαστάσεις μεγαλώνουν, παρουσιάζουμε στα επόμενα σχήματα αναπαραστάσεις της ακολουθίας Halton . Θυμίζουμε ότι η ακολουθία Halton χρησιμοποιεί ένα διαφορετικό πρώτο αριθμό ως βάση σε κάθε διάσταση. Για την πρώτη διάσταση χρησιμοποιεί τον 2, για την δεύτερη τον 3, για την τρίτη τον 5, για την τέταρτη το 7 και ούτω καθ' εξής. Υψηλότερη βάση σημαίνει μεγαλύτερο κύκλο και υπολογιστικό χρόνο. Για την περίπτωση των δύο διαστάσεων, η προβολή της Halton για 1000 πρώτα σημεία είναι η εξής:



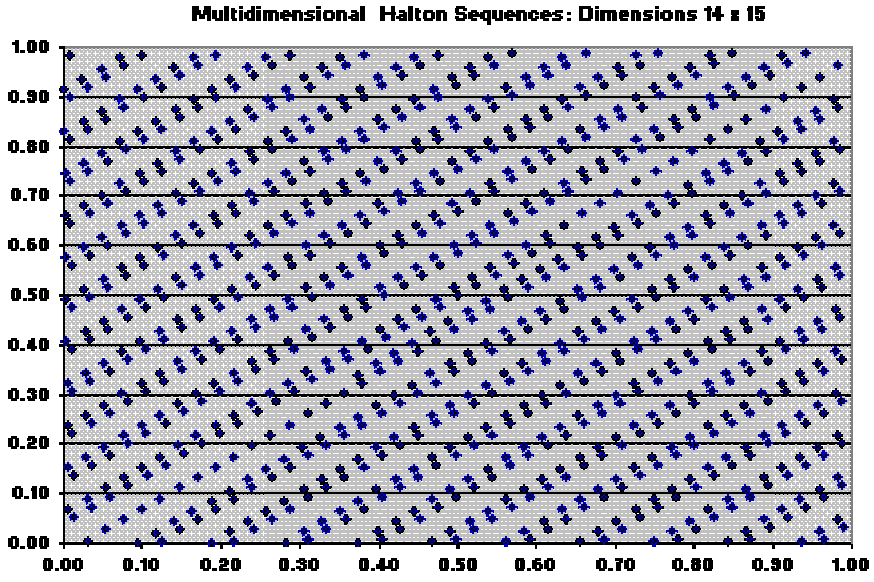
Στο παραπάνω σχήμα, η ακολουθία Halton έχει εκκινήσει από το n=16 και όχι n=1. Αυτό δεν ενοχλεί αφού η ακολουθία διατηρεί τα χαρακτηριστικά της, έτσι ώστε να μην είναι απαραίτητο να ξεκινάμε από το n=0 ή n=1. Επιπλέον υπάρχουν κάποια πλεονεκτήματα αν αποκόπτουμε τους κάποιους n πρώτους όρους της ακολουθίας για να βελτιώνουμε την απόδοση της σε υψηλές διαστάσεις (Galanti & Yung, 1997 pp 69). Η σύγκριση με το αντίστοιχο γράφημα για τους 1000 πρώτους αριθμούς από μια γεννήτρια ψευδο τυχαίων αριθμών (Excel) είναι άμεση.



Σε αυτό το σχήμα φαίνονται καθαρά τα κενά και οι συστοιχισμοί.
Το βασικό πρόβλημα όμως των quasi-τυχαίων ακολουθιών είναι ο εκφυλισμός τους καθώς οι διαστάσεις μεγαλώνουν. Η διαδικασία παραγωγής ομοιόμορφα κατανεμημένων σημείων στον μοναδιαίο υπερκύβο d διαστάσεων γίνεται αυξητικά δύσκολη καθώς το d μεγαλώνει, αφού ο χώρος που πρέπει να καλυφθεί γίνεται πολύ μεγάλος.
Η υψηλής διάστασης ακολουθία Halton εμφανίζει μεγάλα μήκη κύκλων , λόγω της βάσης μεγάλου πρώτου αριθμού. Για παράδειγμα στην διάσταση d=50, χρησιμοποιείται ως βάση ο 50ος πρώτος αριθμός 229, Το μεγάλο μήκος του κύκλου σημαίνει ότι η υψηλής διάστασης ακολουθία χρειάζεται αρκετούς αριθμούς για να καλύψει πλήρως το διάστημα [0,1). Αυτό συνεπάγεται ότι η ταχύτητα παραγωγής των αριθμών που θα καλύπτουν περισσότερο το διάστημα θα μειώνεται καθώς οι διαστάσεις αυξάνονται.
Παρακάτω φαίνεται η ακολουθία Halton για τις 20χ21 διαστάσεις και βλέπουμε ήδη τα κενά και τον εκφυλισμό:



Το παρακάτω σχήμα δείχνει ότι ενώ η ακολουθία φαίνεται ακόμη ομοιόμορφη στις δύο διαστάσεις, εμφανίζει συσχετισμούς μεταξύ των διαστάσεων. Έτσι η ακολουθία Halton γίνεται μη ικανοποιητική μετά από την 14 διάσταση περίπου.



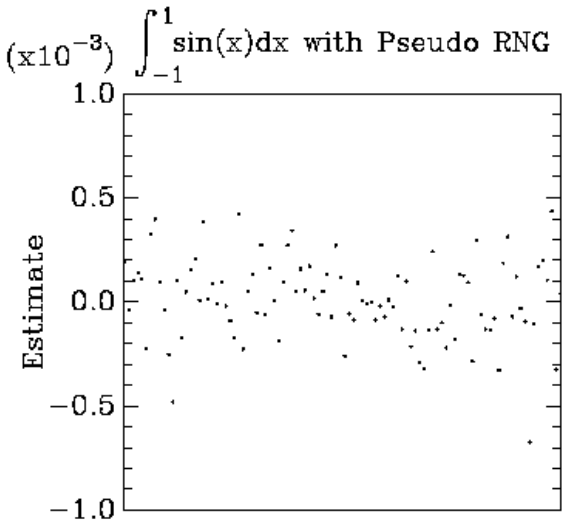
Στην πράξη, αποφεύγουμε να την χρησιμοποιούμε για περισσότερες από 6 – 8 διαστάσεις.

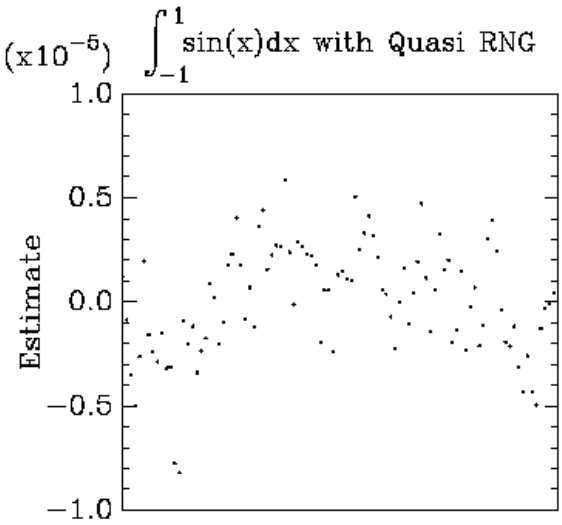
7. Παραδείγματα

Τα παρακάτω παραδείγματα είναι από μια εργασία του Austen McDonald, πάνω στην ολοκλήρωση Monte Carlo (βλ. Βιβλιογραφία 8). Τα σχήματα, όπως αναφέρει, παρήχθησαν από πειράματα Monte Carlo σε ένα Sun Server με 4 επεξεργαστές ταυτόχρονα σε λειτουργία.

Α) Προσέγγιση του ολοκληρώματος $\int_{-1}^{1} \sin(x)dx$

Το ολοκλήρωμα του ημιτόνου από το –1 εώς το 1, εκτιμήθηκε 100 φορές χρησιμοποιώντας τόσο την κλασσική Monte Carlo (με την γεννήτρια ψευδοτυχαίων αριθμών Mersenne Twister 19937 που έχει περίοδο $2^{19937}-1$) όσο και τη quasi Monte Carlo με μια quasi-ακολουθία Sobol' , με αριθμό σημείων n=400.000. Τα παρακάτω σχήματα είναι σαφή και δείχνουν ότι για τον ίδιο αριθμό σημείων n, η ακρίβεια της quasi-MC είναι περίπου δύο τάξεις μεγέθους καλύτερη.

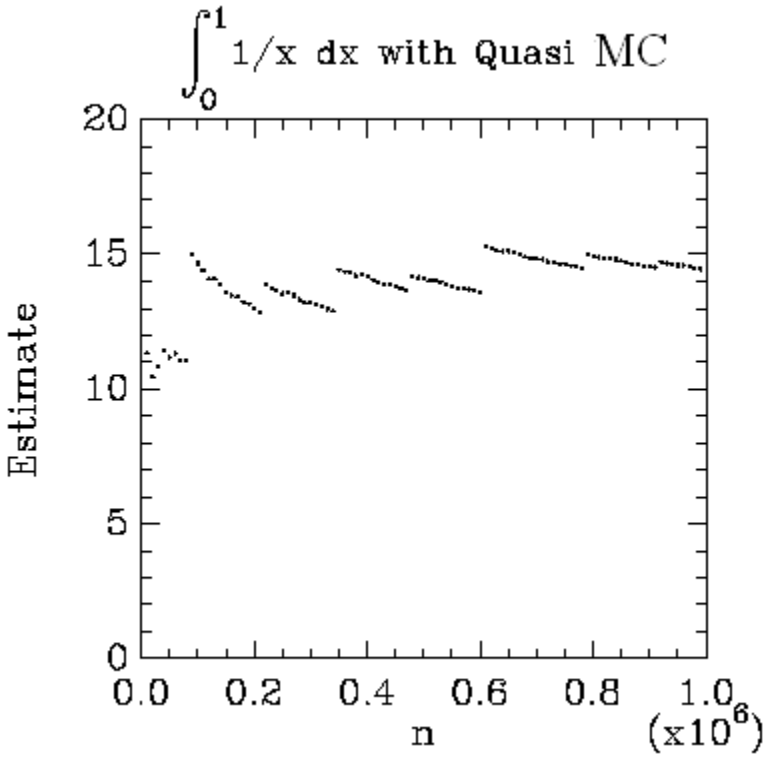




Β) Προσέγγιση του ολοκληρώματος $\int_{0}^{1} \frac{1}{x} dx$

Το παράδειγμα αυτό είναι παραστατικό για την αξία των πληροφοριών που (πρέπει να) έχουμε για την προς ολοκλήρωση συνάρτηση. Είναι σημαντικό να γνωρίζουμε που η συνάρτηση είναι μη φραγμένη, δηλαδή που τείνει στο άπειρο. Σε αυτές τις περιοχές η Monte Carlo, συνήθως δίνει ανακριβείς εκτιμήσεις, αφού δεν υπάρχει φράγμα της συνάρτησης γύρω από αυτά τα σημεία. Στο παράδειγμά μας έχουμε 1/x → άπειρο, καθώς x→ 0 . Στο παρακάτω σχήμα φαίνονται τα αποτελέσματα που μας έδωσε η quasi-Monte Carlo με ακολουθίες

Sobol'. Παρατηρούμε ότι η προσέγγιση δεν διακυμαίνεται πολύ για δεδομένο n, ακόμη και αν ανεβάσουμε τον πλήθος των δειγμάτων από 40000 σε 100 εκατομμύρια. Είναι προφανές ότι είναι μια τελείως ανακριβής εκτίμηση.



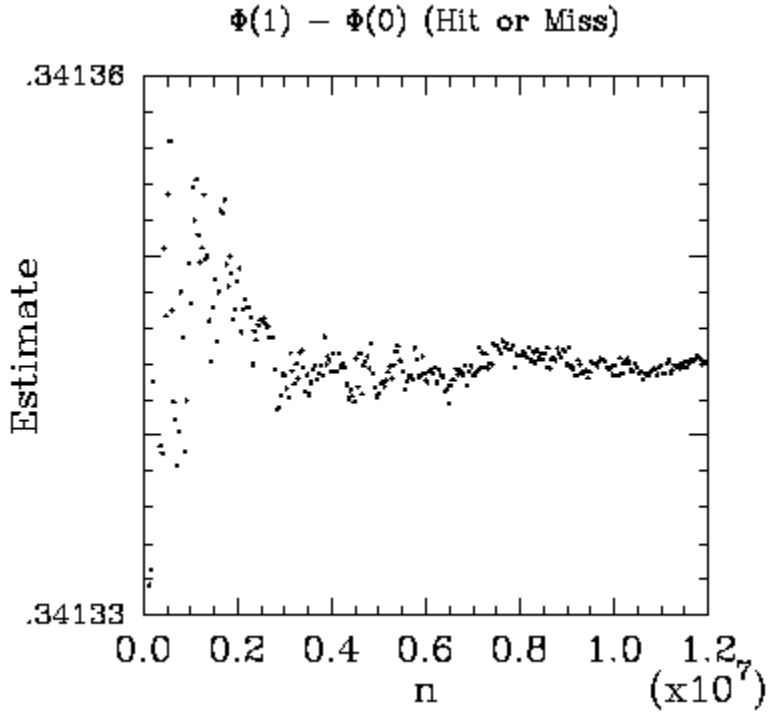
Γ) Υπολογισμός του ολοκληρώματος $I = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$

Το ολοκλήρωμα αυτό εκφράζει στην ουσία (Κ.Ο.Θ) τη πιθανότητα μια stantard κανονική μεταβλητή να πάρει τιμές στο διάστημα [α,b]. Το ολοκλήρωμα αυτό δεν υπολογίζεται εύκολα γιατι δεν έχει απλή ολοκλήρουσα. Συνήθως, χρησιμοποιούμε πίνακες.
Εκτιμούμε το ολοκλήρωμα αυτό χρησιμοποιώντας Hit or Miss και κλασσική Monte Carlo. Θεωρούμε α=0, b=1, με πλήθος σημείων n να κυμαίνεται από 40.000 εώς 12 εκατομμύρια με βήματα των 40.000. Η πραγματική απάντηση σε 12 σημαντικά ψηφία είναι: 0.341344746069.

Αποτυχία ή Αστοχία (Hit or Miss)

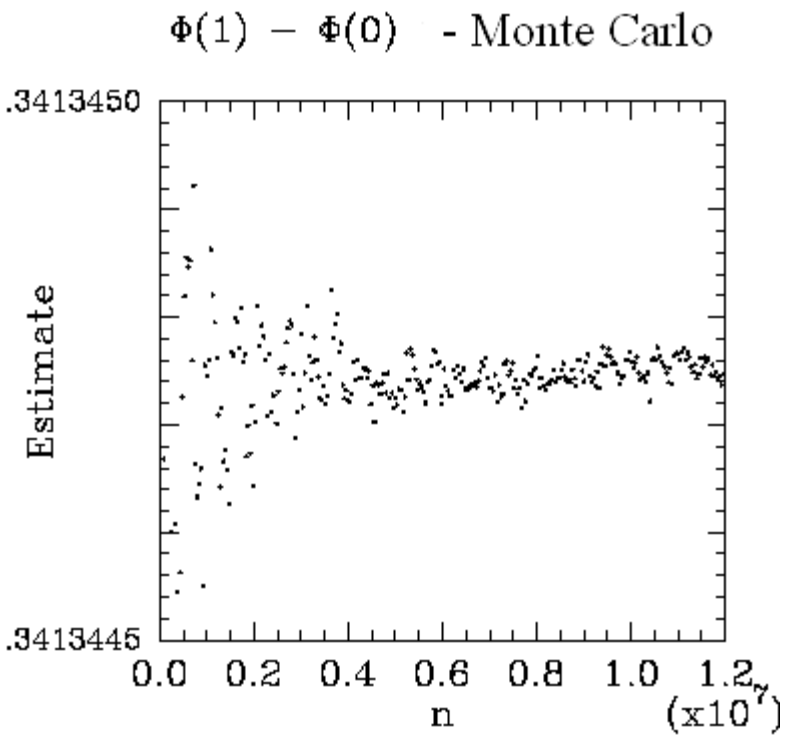
Η μέθοδος αυτή όπως φαίνεται από το σχήμα έχει περιορισμένη αλλά ακόμη αποδεκτή ακρίβεια. Για την τιμή n=40000 επέστρεψε μια εκτίμηση ακριβή σε

τέσσερις δεκαδικές θέσεις. Με το τέλος του πειράματος η ακρίβεια είχε ανέβει στις έξι δεκαδικές θέσεις.

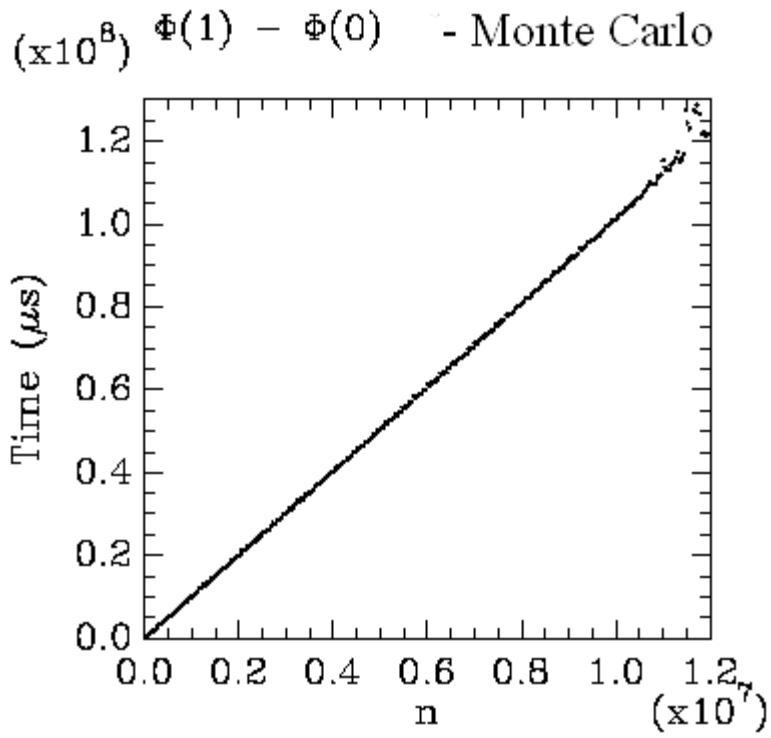


Κλασσική Monte Carlo

Η διαφορά της Monte Carlo με την Hit or Miss είναι ξεκάθαρη. Η εκτίμηση που μας δίνει με 40000 σημεία είναι ακριβής σε 5 σημαντικά ψηφία, όσα δηλαδή η Hit or Miss με 1.3 εκατομμύρια σημεία! Με 20 εκατομμύρια ψηφία, η MC έχει φτάσει στις εννιά δεκαδικές θέσεις ακρίβειας.



Μία ακόμη ενδιαφέρουσα αλλά αρκετά αναμενόμενη παρατήρηση είναι και στις δύο μεθόδους που χρησιμοποιήθηκαν η αύξηση της ακρίβειας ήταν ασυμπτωτική. Δηλαδή, όσο το n ήταν μικρό, μία μικρή αύξηση του προκαλούσε μια σχετικά μεγάλη αύξηση στην ακρίβεια. Όμως, όταν το n είχε γίνει ήδη αρκετά μεγάλο, μια σημαντική αύξηση στην ακρίβεια απαιτούσε ένα πολύ μεγάλο «άλμα» στον αριθμό των τυχαίων τιμών n. Για παράδειγμα, στην κλασσική Monte Carlo, για τις τιμές του n μεταξύ των 6 εκαταμμυριών και των 12 εκατομμυριών υπάρχει πολύ μικρή αύξηση στην ακρίβεια αν και πρακτικά ο αριθμός των τυχαίων αριθμών που χρησιμοποιούνται διπλασιάζεται.
Το παρακάτω διάγραμμα δείχνει τη σχέση χρόνου-αριθμού δειγμάτων που χρησιμοποιούνται και προφανώς προκύπτει μια γραμμική σχέση:



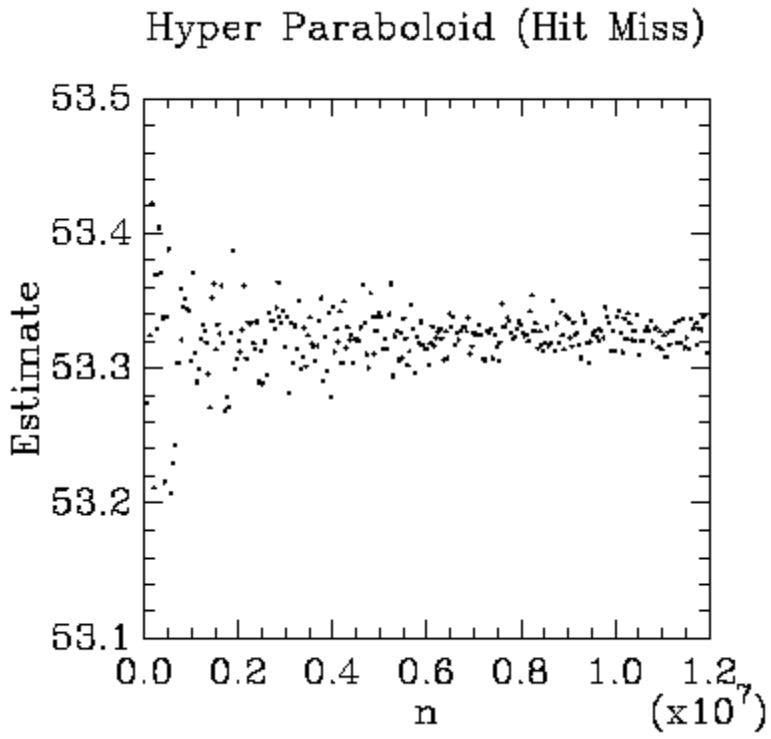
Δ) Πενταδιάστατο Υπερ-Παραβολοειδές

Θα ήταν παράλογο όντως να προσπαθούσαμε να υπολογίσουμε το παρακάτω ολοκλήρωμα του πενταδιάστατου υπερ-παραβολοειδούς:

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1}\left(x_1^{\,2}+x_2^{\,2}+x_3^2+x_4^2+x_5^2\right)dx_1dx_2dx_3dx_4dx_5$$

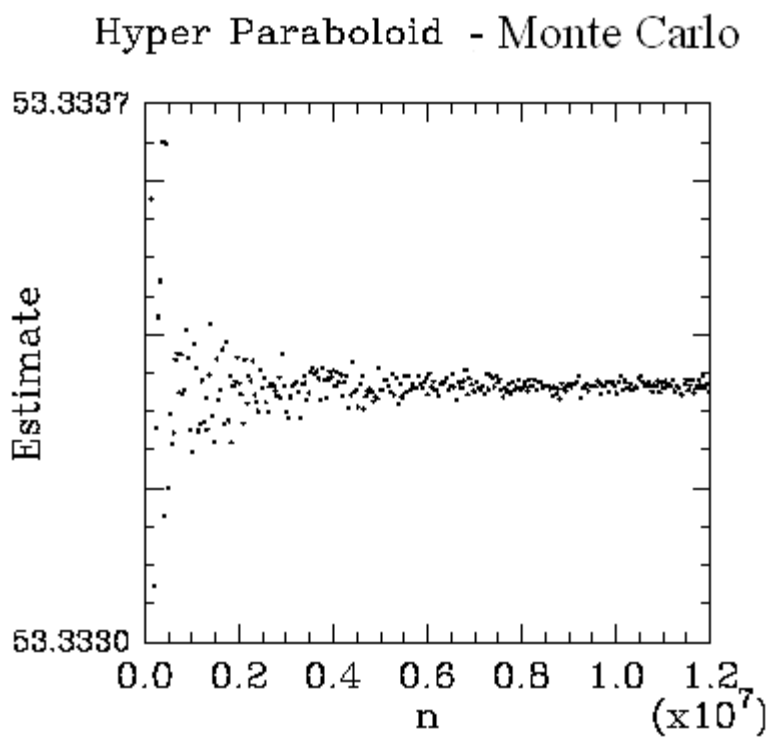
με κάποια αριθμητική μέθοδο πλέγματος ή με την Ριμάνεια μέθοδο αθροίσματος.
Καταρχάς, λύνεται με αναλυτικό τρόπο εύκολα και είναι 160/3=53.3333…. .
Ας δούμε όμως η Monte Carlo , πόσο ακριβής είναι γι' αυτό το ολοκλήρωμα.

Επιτυχία ή Αστοχία
Με την επιτυχία ή αστοχία (Hit or Miss Method) φτάνουμε σε ένα παρόμοιο με το προηγούμενο παράδειγμα συμπέρασμα. Η εκτίμηση δεν είναι τόσο καλή όσο θα μπορούσε να είναι. Καταρχάς φαίνεται να υπάρχει μεγάλη διασπορά στις εκτιμήσεις για μικρό n. Για n=40000, η εκτίμηση είναι 53.274 και διακυμαίνεται μεταξύ 53.2 και του 53.5 μέχρι ο αριθμός των τυχαίων σημείων να γίνει n=800.000. Μια λογική εξήγηση γι' αυτό πρέπει να είναι η μεγάλη διάσταση. Από το σημείο αυτό μέχρι και το τέλος του πειράματος πετυχαίνουμε μόνο μια επιπλέον δεκαδική θέση. Φαίνεται λοιπόν ότι η άποψη που διατυπώσαμε στην αρχή πως η Αποτυχία ή Αστοχία είναι αργή και μικρής ακρίβειας είναι αληθής.

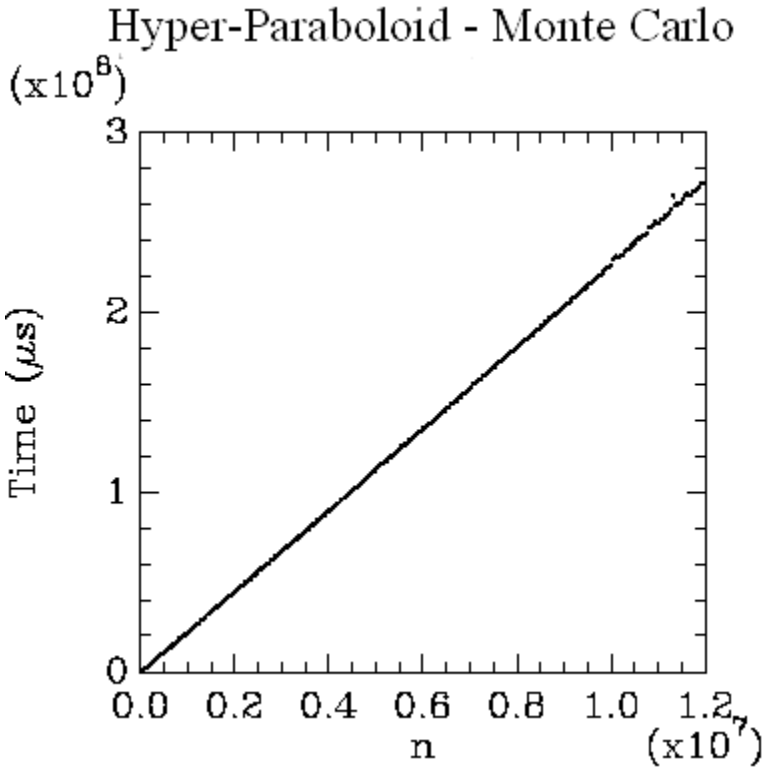


Κλασσική Monte Carlo

Για n=40.000 παίρνουμε ένα αποτέλεσμα ακριβές για 4 σημαντικά ψηφία, δηλαδή για δύο δεκαδικές θέσεις. Παρατηρούμε ότι η ακρίβεια είναι μικρότερη από ότι ήταν στο προηγούμενο παράδειγμα της μονοδιάστατης περίπτωσης. Μπορούμε να υποθέσουμε λογικά ότι αυτό οφείλεται στις περισσότερες διαστάσεις. Επίσης, η ακρίβεια εμφανίζει την ίδια ασυμπτωτική συμπεριφορά όπως και πριν.



Το παρακάτω γράφημα εμφανίζει την σχέση χρόνου-αριθμού δειγμάτων :



Παρατηρούμε ότι ο χρόνος είναι γραμμικός, μόνο που έχει αυξηθεί δραστικά. Αυτό σίγουρα οφείλεται στο ότι παράγουμε πολύ περισσότερα σημεία στην πολυδιάστατη περίπτωση –άν έχουμε n σημεία στις m διαστάσεις κάθε σημείο έχει m συνιστώσες, συνολικά παράγουμε nm σημεία. Προφανώς όμως αυτή η αύξηση δεν είναι τίποτε μπροστά στο χρόνο που θα μας έπαιρνε για να υπολογίζαμε το ολοκλήρωμα με κάποια κλασσική αριθμητική μέθοδο, όπως με το Ριμάνειο άθροισμα.

Από τα τελευταία παραδείγματα επιβεβαιώθηκε η υπόθεση μας ότι η κλασσική ή δειγματικού μέσου (sample mean) Monte Carlo είναι πολύ καλύτερη από την «Αστοχία ή Επιτυχία» . Επίσης, είδαμε στην πράξη την ασυμπτωτική συμπεριφορά των εκτιμήσεων αλλά και την γραμμική εξάρτηση του χρόνου υπολογισμού από τον αριθμό των δειγμάτων n. Τέλος, φάνηκε, αν και δεν υπάρχουν σαφείς ενδείξεις για επιχειρήματα, ότι η Monte Carlo έχει και αυτή κάποιο, αν και αρκετά μικρό, πρόβλημα με την διαστατικότητα.

8. Βιβλιογραφία

1) “Monte Carlo & quasi-Monte Carlo Methods”, Russel E. Caflisch, ACTA Numerica (1998), pp 1-49
2) “A Primer In Monte Carlo”, Ilya Sobol’, Russia
3) "Introduction to Monte Carlo Methods", STEFAN WEINZIERL, NIKHEF Theory Group, Netherlands (PDF)
4) “Lecture Notes In Monte Carlo Integration”, Matias Eklof, dpt of Economics , Uppsala University, 2001 (PDF)
5) “Notes On quasi-Monte Carlo”, Robert Almgren (PDF)
6) “Monte Carlo, quasi-Monte Carlo, and hybrid-Monte Carlo Methods in Computational Finance: Their Current Use And Possible Future Directions”, Giray Okten, CRIAMS, Claremont Graduate University & dpt of Mathematics, Ball State University (PDF)
7) “Random Number Generation and Monte Carlo Methods”, J.E. Gentle, 1998
8) “Monte Carlo Integration”, Austen McDonald, austen@cc.gatech.edu (σχήματα κεφαλαίου 7)

9) Δικτιακός Τόπος από όπου αντλήθηκαν τα γραφήματα για τις quasi – random ακολουθίες του κεφαλαίου 6: www.puc-rio.br/marco.ind/monte-carlo.html